NAUMAN III, Louis W., 1927-
THE PRIMARY STRUCTURE OF THE MYOGLOBINS OF SIX SPECIES OF THE ORDER PINNIPEDIA: A COMPARATIVE STUDY.

University of Alaska, Ph.D., 1971
Biology

University Microfilms, A XEROX Company, Ann Arbor, Michigan

THE PRIMARY STRUCTURE OF THE MYOGLOBINS OF SIX SPECIES OF THE ORDER PINNIPEDIA: A COMPARATIVE STUDY

A

DISSERTATION

Presented to the Faculty of the
University of Alaska in Partial Fulfillment
of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

by

Louis W. Nauman, B.S., M.S.
College, Alaska
May 1971

THE PRIMARY STRUCTURE OF THE MYOGLOBINS OF

SIX SPECIES OF THE ORDER PINNIPEDIA:

A COMPARATIVE STUDY

APPROVED

E. M. Scott

Betty Anne Hudson

Warren Smith

Laurence Irving

Chairman

Vice President for Research and Advanced Study

13 Aug 70

Date

ABSTRACT

Tryptic peptide maps were prepared for the myoglobins of harbor seal, ribbon seal, bearded seal, Weddell seal, northern fur seal, and walrus. The maps were produced by high voltage electrophoresis at pH 1.9, using 20 cm X 30 cm rectangles of Whatman 3MM filter paper. All but 14 residues of the myoglobins of harbor seal and ribbon seal were recognized. The composition of tryptic peptides isolated was identical to the amino acid composition given for the tryptic peptides of Maine coast harbor seal myoglobin.

The myoglobins of four phocid seals (harbor, ribbon, bearded, Weddell) were identical in all peptides recovered (125 of 153 residues). The myoglobin of the northern fur seal exhibited 8 amino acid replacements when compared with harbor seal myoglobin (97 of 153 residues). One difference was seen between myoglobin of fur seal and myoglobin of walrus (90 of 153 residues).

The identity of the myoglobin of the phocid seals studied indicates that they are very closely related. The differences seen between the phocid myoglobin and that of the walrus and northern fur seal is consistent with current systematics.

A model for the treatment of amino acid differences is proposed and applied. This model is based on the hydrogen bonding propensity of the amino acid side chains and the "hydrophilic-out, and hydrophobic-in" principle.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF ILLUSTRATIONS

INTRODUCTION

A protein may be defined as a naturally occurring polymer whose fundamental subunits are amino acids. These amino acids are linked together in a linear sequential fashion by a bond between the terminal carboxyl group of one and the alpha amino group of the next. This bond which forms proteins is termed the peptide bond. The natural production of a protein, qualitatively and quantitatively, is genetically determined.

There are twenty commonly occurring amino acids found in proteins, and the exact sequence in which they occur in a given protein is determined by the genetic material controlling the synthesis of the protein under consideration. It follows then, that the exact sequence of a given protein maybe characteristic for the species in which the protein is produced. Such an intuitive conclusion is verifiable in the data presently at hand. While it is true that only a few proteins from a small number of species have been examined, the results indicate that, by and large, the primary structure of a given protein can be taken as a character for a particular species.

There is, however, a good deal of similarity in the primary structure of proteins from closely related species. In fact, the sequences are so similar that it is the differences that have been the basis for comparison rather than the similarities. The two proteins that have received the most effort are cytochrome c and hemoglobin. While the information resulting from these studies

may be considered from a number of viewpoints, the most immediate one is that of evolution. Species differences in a given protein may be explained on the same basis that one explains other species differences. If one accepts the occurrence of random mutations in the chromosome then a change may be expected in the end product of the chromosome, namely the protein primary structure. The persistence of such a change, an amino acid replacement, should depend on the impact of that replacement on the molecule, its structure and its action. The degree to which a given replacement persists or becomes established in a given population or species will depend to some degree upon this impact and upon the circumstances surrounding that population or species. At first thought, it may seem impossible for a single mutation to become characteristic of a population of thousands or even millions but cataclysmic declines in population may leave the carrier of an atypical protein sequence at a breeding advantage with the result that the protein sequence that is now imprinted on the population or species is not really due to selective advantage. Thus, for example, the bison fell from a population of 50,000,000 to perhaps 50 before the decline was reversed. It should also be noted that many, if not most, of the amino acid replacements seen are probably trivial or neutral, and should be viewed as reflecting a genetic change without selective advantage.

It is true then, that various species many times differ in the primary structure of the same protein. However, there appears to be different degrees of constraint under which different proteins vary.

This degree of constraint should relate to the function of the molecule and its importance to the organism. The bovine milk proteins, for instance, are known to vary extensively within a given breed, (Peterson et al. 1966). This is acceptable functionally because the primary role of milk proteins is to provide protein to the young mammal. Cytochrome c, functioning near the end of the transport chain as an electron transport molecule, possesses the same sequence in both man and the chimpanzee, species belonging to the same order but in different families.

This example, one of many that could be cited, indicates that from the standpoint of variability, proteins are not equivalent. Nevertheless, from the molecular point of view, the greater the similarity between the primary structure of the same protein from different taxonomic groups, the more closely these organisms are related genetically. It is of interest to correlate established relationships with similarities or differences in the primary structure of the proteins.

The evidence to date, in relation to the almost infinite array of biological species, is too fragmentary to confirm general principles. However, a few workers in the field have made an impressive beginning. A major effort has centered around cytochrome c. The primary structure of the cytochrome c of twenty species, from lower plants to man, have been determined. On the basis of this information, phylogenetic trees have been constructed (Nolan and Margoliash, 1968). The resulting schema is consistent with current systematics.

The terms homology and homologous have been applied to proteins in much the same way that they are commonly used for morphological features in biology. Margoliash (1969) stated that the term homologous, as applied to proteins, implies "that the genetic coding for the polypeptides considered, in all the species carrying these proteins, had at one time a common ancestral gene".

Cytochrome c

Cytochrome c is the only protein to date which has been studied in both plant and animal kingdoms. Dickerson and Geis (1969) have estimated the number of amino acid replacements in cytochrome c between major groups. This information, reproduced in part as Table 1, shows that the greater the separation between taxonomic groups, the greater the number of amino acid replacements.

If the primary protein structure of distantly related species differs to the degree seen, how closely must species be related to show no difference? As stated previously, cytochrome c from man and from chimpanzee representing different families within one superfamily have the same sequence. Man versus rhesus monkey, representing different superfamilies shows a single difference within the same suborder. The sequences of cytochromes c of the pig and that of the cow representing different suborders (order Artiodactyla) are identical. The cytochrome c from duck (order Anseriformes) possesses three differences when compared to the sequence of cytochrome c of the pigeon (order Charadriiformes). This suggests that the taxa are not equivalently separated; there is not a single constant number of

Table 1

Amino acid substitutions between
the cytochromes c of various species.
(Data is from Dickerson and Geis (1969)

Groups Compared*	Average number of** substitutions
Mammals (11) vs. birds (4)	10.0
Mammals (11) vs. reptiles (3)	14.3
Mammals (11) vs. fish (1)	18.5
Mammals (11) vs. invertebrates (2)	25.4
Mammals (11) vs. plants (4)	46.7
Higher vertebrates (18) vs. fish (1)	18.5
Vertebrates (19) vs. invertebrates (2)	25.9
Animals (21) vs. plants (4)	47.0

* The numbers in parenthesis refer to the number of species in the group described.

** Average of differences between all pairs of species.

primary structure differences for a given protein between taxonomic groups, i.e. that substitution may be more effectively incorporated into some groups than into others. Additionally, the time that the taxonomic groups have been separated from one another certainly is a factor.

Hemoglobin

The degree of constraint seen in cytochrome c is in contrast to that observed in hemoglobin. Normal hemoglobin is composed of two different chains termed alpha and beta, and both chains have been shown to vary when hemoglobins of organisms from rather closely related groups are compared. Nolan and Margoliash(1968) note that primate hemoglobins furnish the most extensive data for comparison of primary protein structure between closely related species. The complete sequence for both chains in the hemoglobins from man, chimpanzee, and gorilla is known. Human and chimpanzee hemoglobins are identical and differ from gorilla hemoglobin by a single replacement in each chain. As taxonomic separation increases, so does the number of observed differences. In addition to the three primates listed above, the sequence of the alpha chain of the rhesus monkey hemoglobin is also known. There are four differences between the alpha chain of rhesus monkey hemoglobin and that of human and chimpanzee. A similar comparison of the alpha chains of rhesus monkey and gorilla indicates five differences. The beta chains from the hemoglobins of five primates are known. Comparisons of the alpha and beta chains similar to those already mentioned are listed in Table 2. Within the order of

Table 2

Comparison of alpha and beta chains
of some primate hemoglobins

Groups Compared	Differences α	Differences β
Man (hominidae) vs. chimpanzee (pongidae)	0	0
Man (hominidae) vs. gorilla (pongidae)	1	1
Hominoidae (3) vs. spider monkey (ceboidae)		5
Hominoidae (3) vs. rhesus monkey (cercopithecoidae)	4	8
Ceboidea (1) vs. Cercopithecoidae (1)		8

primates, then, as many as 12 amino acid replacement differences in hemoglobin have been noted between two species (man versus rhesus monkey).

Hemoglobin structure can vary within a single species. In human hemoglobins there are distinct differences within the species. Human hemoglobin S is perhaps the best known example of this variation. Here the "normal" glutamic acid at position #6 is replaced by valine. This is a special case which apparently became established in Africa due to the resistance to the malarial parasite conferred upon the heterozygous individual. There are many other atypical hemoglobins that do not seem to be associated with any distinct advantage or pathology. The 1969 Atlas of Protein Sequence and Structure (Dayhoff, 1969) lists 101 amino acid replacements in human hemoglobin. For the most part, the replacements are limited to single substitutions but a few examples of more than one replacement have been observed. In hemoglobin C Harlem there are two replacements in the beta chain, and in hemoglobin X there is one replacement in the alpha chain and one replacement in the beta chain.

In addition to replacements, deletions have also been found. In hemoglobin Freiburg the valine at position #23 in the normal beta chain has been deleted and in hemoglobin Koelliker, the arginine present at position #141 in the normal alpha chain has been deleted. Hemoglobin Gun Hill is an example of a hemoglobin in which a deletion of several amino acids has occurred at positions

#91-95 of the normal beta chain.

According to Sochard (1968) ten variant cases were found in a study of hemoglobins of 8,000 Europeans, about 1:1000. In a similar study of 300,000 Japanese, only 26 abnormal cases were found, or about 1:10,000. The European continent has been subjected to successive invasions by many different races, while the islands of Japan have been isolated and insulated from such genetic diversity. The observed number of variants must represent a minimum number since most of them have been detected by differences in electrophoretic migration. As Vogel (1969) noted, 82 out of 91 amino acid replacements observed in α, β, γ, and δ chains involved amino acid pairs with different charges. Since 15 of the protein amino acids are neutral and only 3 are basic and 2 acidic there is the expectation that many more replacements do exist, but since they do not involve differences in charge they are not detected by electrophoresis.

The appearance of occasional variants or mutants is, of course, consistent with, and indeed essential to, the general theory of evolution. A change must exist before it can become fixed in a population or species. The persistence of a given change is the result of the sum of all the interactions between the total molecule and its environment, not only at the cellular and organismic level but also at the level of that organism's ecological niche. There is then a well-founded expectation for the appearance of occasional variants of that protein primary structure which is typical of the taxonomic group under consideration. The hemoglobin variants that

are found may be taken to be exceptional then, and not typical of the species.

Myoglobin

While cytochrome c and hemoglobin have been the object of extensive study, the other major heme protein, myoglobin, has received less attention in spite of the fact that both the three-dimensional structure and the primary structure have been established for the myoglobin of the sperm whale.

The complete structure of the myoglobins of sperm whale, harbor seal, porpoise, and horse is known, and the sequence of human myoglobin has been almost completely determined. The partial sequence of myoglobin from humpback whale and cow has been determined.

Sperm Whale - (*Physeter catodon*)

Edmundson and Hirs (1962, a,b,c) and Edmundson (1963, 1965) established the primary structure of sperm whale myoglobin. Column chromatography of the tryptic peptides and the chymotryptic peptides was the basic approach. Sequential Edman degradation provided the sequence in each peptide. Overlapping peptides, peptides containing the amino terminal sequence of one peptide and the carboxyl terminal sequence of another, were obtained as a result of the use of the two different enzymes.

Harbor Seal - (*Phoca vitulina concolor*) and Porpoise - (*Phocaena phocaena*)

Bradshaw and Gurd (1969) published the complete sequence for both of these species in a series of five papers. Again the initial

separation of the peptides was performed by means of column chromatography with final purification by electrophoresis or paper chromatography as required. There are twenty-five amino acid differences between sperm whale myoglobin and harbor seal myoglobin. Fifteen differences are noted between the myoglobins of the sperm whale and the porpoise. Comparison of harbor seal myoglobin with porpoise myoglobin shows that they differ by thirteen amino acid replacements.

Horse

Dautrevaux et al. (1969) established the primary structure of horse myoglobin. Comparison of horse myoglobin with myoglobin of the sperm whale shows nineteen differences.

Human

Hill et al. (1969) have determined the major portion of the sequence of human myoglobin. The portion remaining is not entirely unknown but the authors were not able to determine a few short sequences in a totally unambiguous fashion. There is a deletion in human myoglobin corresponding to position #9 in the other myoglobins so that human myoglobin possesses 152 residues rather than the 153 found in the other species.

Partial sequences of myoglobins of other species have been reported.

Humpback Whale - (*Megaptera novaeangliae*)

Edman and Begg (1967) reported the first sixty residues (from the amino terminus) of humpback whale myoglobin. There are six differences between the myoglobins of the sperm whale and that of the humpback whale. The sequence was determined by means of the

"sequenator" developed by Edman and Begg.

Bovine

The partial sequence of myoglobin of bovine origin has recently been published by Dautrevaux et al. (1969). The peptide sequences missing are in large part those of the "core" encountered by Edmundson and Hirs (1962b). Column chromatography provided the basic separation of the peptide mixture.

Other Species

Stockell (1961) compared the "fingerprints" of the soluble peptides of tryptic digests of myoglobins of sperm whale, horse, harbor seal, porpoise, sei whale, gentoo penguin, and tortoise. The conclusions drawn were based upon peptide spot position and reaction to specific color reagents for some of the amino acids. From the vantage point of some ten years later, this work appears almost inconsequential, but it did help to indicate that the fingerprint method could be of value in comparing some proteins other than hemoglobin.

Jegorov et al. (1965) used a modification of the fingerprinting method for a comparative study of the primary structure of myoglobin from two species of monkey (*Macacus rhesus* and *Macacus cynnamolgus*), horse, dolphin, and sperm whale. The peptide maps of the monkey myoglobins were judged to be identical. The number of replacements are indicated, although the amino acid analyses were, in most cases, qualitative.

Statement of the Problem

It is evident that there is general correspondence between the

number of differences between homologous proteins and their taxonomic separation. A question of interest is whether conventional differences between taxa (genera within a family, families within a superfamily, or order, etc.) may not correspond to a more or less constant number of amino acid replacements in homologous proteins. However, evidence already at hand shows that different proteins behave differently in this regard. Thus, the primary structure of cytochrome c exhibits one difference between man and rhesus monkey, while the beta chain of hemoglobin exhibits eight differences between man and the rhesus monkey. This represents 1 change in 104 residues for cytochrome c or .01 changes per residue, 4 changes in 152 residues or .026 changes per residue for the alpha chain of hemoglobin, and 8 changes in 146 residues for the beta chain of hemoglobin, or .05 changes per residue. This suggests that even though cytochrome c and hemoglobin are both respiratory proteins, there are different constraints upon substitutions in the primary structure of these two proteins.

The degree of constraint on myoglobin is unknown since there is little information available on its primary structure and this makes it an attractive subject for study. The molecule is less complicated than hemoglobin, since it is a monomer with a molecular weight of about 18,000. In addition, highly purified preparations can be obtained with relative ease using chromatographic procedures.

This study is concerned with the myoglobins of six members of the pinnipedia, a group of mammals little studied. One practical advantage to studying the pinnipedia is that as marine, and therefore diving

mammals their skeletal muscles are rich in myoglobin, thereby providing a source of material with myoglobin concentration close to 1%. As marine mammals, they are both available and of particular interest in Alaska whose marine mammal fauna constitutes one of its important resources; indeed, the Bering Sea is the richest area in the world in both the species and numbers of marine mammals.

All three families of the pinnipedia were studied. This included four species of the family Phocidae, three of which were from the Bering Sea and the fourth, the Weddell seal, from the antarctic. The walrus as the sole extant species of the Odobenidae represented that family. The third family, the Otariidae, were represented by the northern fur seal. The myoglobins from these species provided interesting comparisons, not only from the standpoint of systematics, but also from that of geographic isolation.

First, comparison of two geographically distinct subspecies of harbor seal was possible. Our subspecies was Phoca vitulina richardi from the Bering Sea while the myoglobin sequence established for harbor seal by Hartzell, et al. (1969) was determined from Phoca vitulina concolor taken off the Maine coast. Second, it was possible to compare myoglobins from different species of the same family occurring in the same region, since the specimens of the harbor seal, ribbon seal, and bearded seal were all taken in the Bering Sea. A third comparison was that of the myoglobins of geographically isolated species from the same family. The Weddell seal is found only in the Antarctic, while Bering Sea phocids are limited to the northern

hemisphere. The fourth comparison possible was that of the myoglobins of the three distinct families within this order.

While the sequences derived from this study are useful for making comparisons based upon current systematics, they are important in their own right in adding to the limited body of knowledge presently available on myoglobins.

The question this research was designed to answer may now be stated. "What is the degree of difference or similarity seen in the primary structure of the myoglobins of four members of the family Phocidae, and what is the degree of difference or similarity seen in the primary structure of the myoglobins of members of the three families of the pinnipedia?"

EXPERIMENTAL

Biological Materials

Muscle samples from harbor seal, ribbon seal, bearded seal, and fur seal were collected by Mr. John Burns of the Alaska Department of Fish and Game and by Mr. John Baust of this Institute during the Alpha Helix Bering Sea Expedition of March 1968. The specimens were shot, and the skeletal muscle samples were taken immediately and frozen. These materials were subsequently brought to the Institute in the frozen state, and were kept frozen at -20°C until needed.

The walrus material (a single heart) was collected by Mr. Vernon Slwooka of Gambel, St. Lawrence Island. The partially frozen heart was shipped by air to Fairbanks, and immediately transferred to a freezer.

The Weddell seal material was collected for Dr. Laurence Irving during a trip to McMurdo Sound in winter of 1968. The specimen was frozen and kept frozen during its transfer to the Institute.

Species Description and Ranges

Nomenclature follows Scheffer (1958).

Harbor seal - Phoca vitulina richardi. This subspecies represents the eastern Bering Sea population of harbor seals which ranges from Herschel Island in the MacKenzie Bay to the eastern Bering Sea, Aleutian Islands, and southward along the coast to northern Baja California and Mexico. As Scheffer points out, the source of the specimen determines the subspecies classification.

Ribbon seal - Histriophoca fasciata. This pinniped inhabitats the

Pacific-Arctic from Point Barrow southward to the tip of the Alaskan Peninsula and from the Alaskan coast to the coast of Siberia.

Bearded seal - Erignathus barbatus nauticus. The pacific bearded seal is the largest (200-300 kg) phocid seal that occurs in Alaskan waters. This subspecies inhabits the waters of the western Arctic and the subarctic Pacific Ocean.

Weddell seal - Leptonychotes weddelli. This seal is an antarctic species and is circumpolar in range. According to Scheffer, there are only occasional reports of this species north of 70° S latitude.

Walrus - Odobenus rosmarus divergens. The pacific subspecies inhabits the Bering and Chukchi Sea at the edge of the polar ice. Its most southern range is considered to be about 58° N in the Bering Sea.

Northern fur seal - Callorhinus ursinus. The northern pacific eared seal ranges from the Pribilof Islands, Alaska west to Robben Island, U.S.S.R. and Sakhalin, U.S.S.R. In winter and spring, the species is dispersed across the northern Pacific Ocean, southern Bering Sea, Sea of Japan, and Sea of Okhotsk as far south as the waters off Honshu, Japan and San Diego, California.

METHODS

Introduction

The early procedures for amino acid analysis were gravimetric in nature and required gram quantitities of protein. Eventually colorimetric procedures were developed for specific amino acids. The Sakaguchi color reagent for arginine and the phenol reagent for tyrosine and tryptophane are typical present-day survivors of a host of

specific colorimetric methods and reagents.

The general reaction of amino acids with oxidizing agents to form volatile products such as ammonia and carbon dioxide, led to the development of manometric methods. The now classical Van Slyke method for the determination of amino nitrogen is based on the release of stoichiometric quantities of gaseous nitrogen produced upon reaction of α amino acids with nitrous acid.

Microbiological methods for the determination of amino acids in a protein hydrolysate were surprisingly accurate under proper control and permitted the determination of all the amino acids present in the acid hydrolysate of milligram quantities of protein (Snell, 1945). Eventually isotope dilution was added to the list of procedures available to the protein chemist. By the time of the appearance of chromatographic procedures, many laboratories were engaged in the amino acid analyses of proteins by a combination of methods and the list of analyzed proteins was beginning to grow (Tristram, 1949).

The advent of a single method for the simultaneous separation and estimation of the amino acids in the acid hydrolysis of a few milligrams of protein (Moore & Stein, 1951) ushered in a new era for the field of protein chemistry. The original manual procedure required two and one-half days for a single analysis. The automated procedures (Moore et al., 1958; Spackman et al., 1958) reduced this time to one day. The microbiological, gravimetric, colorimetric methods had required, in many cases, the full time of three to four chemists for a year; but now it was possible to complete a compositional analysis in

a matter of days.

With such an analytical tool capable of providing useful data at the level of .02 micromole to .2 micromole, the method of fingerprinting is possible. This reduction in time and effort to the present day level permits the kind of work pursued here. The refinement of the mapping techniques based on amino acid analysis plus the final analyses of the peptides eluted from twenty or so maps, required hundreds of analyses. This number of analyses is possible only because of the present state of the technique.

The peptide mapping procedure employed during this study is an adaptation of the basic "fingerprinting" technique of Ingram (1958). Ingram's method combines paper electrophoresis and paper chromatography for the separation of the mixture of the peptides produced by the digestion of a protein by the proteolytic enzyme trypsin. In this study, an aluminum horizontal water cooled flat plate high voltage electrophoresis cell was used. This permitted the use of voltage gradients up to 150 v/cm to be used to effect the separation in the electrophoretic mode.

The amino acid analyses were performed with an automatic amino acid analyzer built from the basic design of Spackman et al. (1958). The principal modifications incorporated were the use of the single column gradient system of Piez and Morris (1960) and the use of the Technicon roller type pump, which segments the column effluent after injection of ninhydrin reagent[1]. The segmentation of the effluent

[1] The ninhydrin, hydrindantin and methyl cellosolve used for preparation of the ninhydrin reagent were obtained from Pierce Chemical Company.

stream with discrete nitrogen bubbles preserves the integrity of the amino acid peaks as they are eluted from the ion exchange column. The segmented stream passes into a glass coil immersed in bath oil wax at 95°C. The reaction mixture emerges from this coil some 12-15 minutes later. By this time the reaction has proceeded to virtual completion. The reaction vessel effluent passes through a short T which vents the entrapped segmentation bubble. The debubbled liquid passes continuously through colorimeters which are periodically monitored in an automatic fashion to produce a continuous record by way of a three point recorder. This continuous record of the optical density of the reaction stream is the raw data which provides the information the peptide compositions are based upon. Each analysis required six and a half hours. Five self-contained chromotographic systems were programmed by sequential timers to provide continuous operation.

The data was normally grouped on the basis of a ninhydrin batch of nineteen liters, (70-80 analyses). Elution buffers were made up in quantities sufficient to prevent any need for change during the consumption of a given batch of ninhydrin reagent. Color yield factors were newly determined for each of the reagent units.

Separation of Myoglobin

Myoglobin was extracted from muscle and further purified by the procedure given below.

Twenty-five grams of frozen tissue were allowed to thaw, minced with scissors, and extracted twice with an equal volume of water for two minutes in a Virtis homogenizer. The resulting suspensions were

centrifuged and the supernatant was stored frozen until needed.

Chromatographic purifications utilized both diethylaminoethyl-cellulose (DEAE)[2] and SE Sephadex[3]. DEAE cellulose chromatography was carried out with a column 2.5 cm x 30 cm in .001 M Tris[4] buffer, using 4 psi air pressure to pack the column. After overnight equilibration of the column, the untreated extract was added directly to the surface of the absorbent bed.

Ten to twenty milliliters of extract were applied to the column. The column was then eluted with the equilibrating buffer at the rate of 20-25 ml an hour. The cytochrome c was not retained by the DEAE cellulose and moved off the column quickly. The myoglobin fraction was only slightly retarded and was not completely resolved from the cytochrome c. This did result in the partial resolution of some of the cytochrome c present in the extract. While the amount of cytochrome c recoverable in this fashion was usually less than a milligram, it showed that some cytochrome c is extracted under these conditions since this protein's molecular weight was not sufficiently different to permit its resolution from myoglobin under normal conditions. This means that myoglobins prepared by means of gel filtration alone will be contaminated with cytochrome c.

The contents of the tubes containing the myoglobin fraction were pooled, and the resulting solution applied to an SE Sephadex

(2) Available as Cellex-D from BIO-RAD Laboratories, Richmond, Calif.

(3) Available from Pharmacia, Inc., 800 Centennial Ave. Piscataway, N.J.

(4) All chemicals used were Baker Analyzed Reagent Grade unless otherwise specified.

column after the method of Hapner et al. (1968). The solution was applied to a 3 cm x 30 cm SE Sephadex column equilibrated with .01 M phosphate buffer at pH 6.8. The column was washed with 200 ml of .01 M buffer to remove a colorless protein. The myoglobin was then eluted with .01 M phosphate buffer at pH 7.4. The cytochrome c remaining on the column was readily eluted with .02 M disodium phosphate buffer. As judged by spectral ratios, this material was 90% pure cytochrome c. The only mention of cytochrome c as a contaminant in chromatographed myoglobin preparations is found in Hapner et al. (1968). This is significant, since some investigators prepare myoglobin by a single pass through a Sephadex G-75 column. Experimentally, a myoglobin preparation obtained by such Sephadex treatment was seen to possess cytochrome c when examined by an SE-Sephadex column.

Myoglobin prepared by the method outlined above is now at least 95% homogenous by gel electrophoresis at pH 8.5 and by spectral ratios. In general, such a level of purity is adequate for the work described here, since a 5% impurity should not interfere with the subsequent determinations.

Preparation of Globin from Myoglobin

The acetone-insoluble phosphate salt was removed from the myoglobin solution by passage through a column of Bio-gel P-2[5], using water as an eluent. Potassium ferricyanide was added to the sample as a marker to indicate the degree of resolution from the salt. The desalted myoglobin was added to an acid-acetone solution at 4°C

[5] Available from BIO-RAD laboratories, Richmond, California.

(.2 ml conc. HCl in 100 ml acetone). The resulting globin was permitted to settle out. After decantation, centrifugation, and washing with plain acetone at 4°C, it was centrifuged again. After a second washing, the globin was recovered as a fine powder upon evaporation of the acetone. This method provides for rapid preparation of globin with a normal yield of about fifty to one hundred milligrams. Based on the measured myoglobin content of the seal muscle, this is an overall yield of about 10%. Yields could be improved by adopting an acceptable method of concentrating the myoglobin solutions to minimize losses during the removal of the heme group.

Exploratory Work with Cyanogen Bromide Cleavage

Cyanogen bromide will cleave proteins at those peptide bonds whose carboxyl group is contributed by methionine (Gross & Witkop, 1962). Edmundson (1963) used this chemical cleavage to resolve the insoluble core problem in sperm whale myoglobin. Black and Leaf (1965) reported the successful cleavage of horse myoglobin followed by separation of the peptides using a fingerprinting technique on paper. Since both sperm whale myoglobin and harbor seal myoglobin contain two methionine residues at identical positions, it was thought that cyanogen bromide cleavage would be of material value. The rationale being that each of the three large peptides could be isolated as an intermediate in the degradation of the protein to recognizable tryptic peptides. Such a step would provide a firm basis for the assignment of peptides.

The myoglobins of both sperm whale and harbor seal were

subjected to cyanogen bromide cleavage, and the peptides separated by the method of Black and Leaf, (1965). Cleavage of the globins was achieved and three peptide spots were identified on the maps. One of these spots proved to be the carboxyl terminal peptide, and the analyses were quite clear (Table 3).

However, the peptide mapping of the tryptic digest of the isolated fragment provided only marginal resolution. The peptide map was clear enough but the amounts of peptide recovered for analysis was less than satisfactory. Some of the peptides were at levels sufficient for analysis, while others were not. In addition, the remaining two peptides, one of 55 residues, the other of 72 presented problems both from the standpoint of recognition and analysis. The evaluation of the method indicated that the proper cleavage did occur, but recovery problems were of such a nature that the method had little hope of success. It was therefore decided to adopt the tryptic cleavage of the total molecule followed by fingerprinting.

Tryptic Hydrolysis of Globin

This enzyme is used extensively to cleave proteins into a set of component peptides. The elucidation of the primary structure of ribonuclease by Smyth et al., (1963) was the first major successful use of this enzyme in protein sequence investigations.

Trypsin is usually the first cleavage reagent employed in primary structure work because it is quite specific for those peptide linkages in which lysine or arginine contribute the carboxyl group. The commercial availability of highly purified trypsin

Table 3

Carboxyl terminal peptide from cyanogen bromide cleavage

	Sperm whale			Harbor seal	
	exp.	ref.		exp.	ref.
Asp	2.27	2	Asp	2.19	2
Glu	3.29	3	Glu	2.02	2
Gly	2.08	2	Gly	2.15	2
Ala	2.75	3	Ala	2.77	3
Ile	1.00	1	Ile	.83	1
Leu	2.88	3	Leu	2.66	3
Tyr	1.45	2	Tyr	.91	1
Phe	.94	1	Phe	2.28	2
Lys	4.24	4	Lys	3.68	4
Arg	1.10	1	His	1.21	1
			Arg	1.13	1

preparations plus the high degree of reliability and reproducibility of the results of tryptic digestion are two further factors that have led to the general acceptance of trypsin as a primary reagent in studying protein structure.

One problem that has always been associated with the use of trypsin is contamination with chymotrypsin. Chymotrypsin cleaves those peptide bonds in which the carboxyl group is contributed by such apolar amino acids as leucine and isoleucine. Over the years, many attempts have been made to eliminate this activity from preparations of trypsin with reduction in chymotryptic activity to less than one percent. The preparation used in the work reported here had a specified chymotryptic activity of .24%. This figure is based on an assay with a synthetic substrate activity against its biological substrate.

As will be seen in the experimental results, the trypsin preparation employed in this work cleaved the myoglobin chain at points that do not involve either lysine or arginine. Whether this cleavage is due to an inherent activity in trypsin or is due to a chymotryptic impurity is immaterial. As Rovery (1967) pointed out, the limited cross specificity of these two enzymes may be rooted in the similarities in structure of the active centers of chymotrypsin and trypsin. The fact that these unexpected cleavages did occur in a reproducible fashion actually proved to be a help rather than a hindrance. At any rate, these "atypical" cleavages do not destroy the validity of the use of the trypsin preparation as a peptide

producing reagent.

Procedure for Tryptic Hydrolysis of Globin

The globins were digested with trypsin[6] at an enzyme-substrate weight ratio of 1 to 100, using the following procedure. Fifteen milligrams of the globin were suspended in 1.5 ml of .5 M ammonium bicarbonate (pH 7.8). To this suspension .15 ml of trypsin solution (2 mg in 2 ml of .001 N HCl) was added. The reaction was permitted to proceed at room temperature (27°C) in a stoppered test tube. The pinniped globins were not readily soluble in this buffer but within an hour after the addition of the trypsin, the solution became clear. No subsequent precipitation occurred. An additional .05 ml of the trypsin solution was added eighteen hours later. At the end of twenty-four hours, the solution was divided into three aliquots in 3" test tubes and placed in a vacuum desiccator containing sodium hydroxide pellets and concentrated sulfuric acid. The desiccator was then evacuated to 5 mm of Hg with an aspirator pump. The ammonium bicarbonate was readily removed by this procedure and the peptide mixture was obtained as a tannish-yellow glassy material.

Procedure for Peptide Mapping

Whatman 3MM filter paper was used as the support medium for the peptide maps. Strips, 55 cm long were cut from a 27 cm wide roll and washed overnight with 10% (v/v) formic acid by descending chromatography. The washed strips were allowed to dry in a fume hood at

(6) Crystalline bovine trypsin purchased from Gallard-Schlesinger Chemical Mfg. Corp. 584 Mineola Ave. Carle Place, Long Island, .N.Y.

room temperature for at least eighteen hours, since failure to remove traces of the acid results in excessively high current flow during electrophoresis.

Disposable plastic gloves were worn during all subsequent manipulations of the paper, and all work surfaces were covered with fresh filter paper. These precautions are necessary to minimize contamination with ninhydrin positive material.

Electrophoresis wicks (10 cm x 32 cm) were cut from Miracloth[7] (Burns and Turner, 1967). Dialysis tubing was slit along one edge to produce barriers to electroosmosis (Bieleski, 1965). The dialysis tubing and the wick material were soaked in buffer. Prior to use, they were rinsed with a fresh volume of buffer.

The wicks were placed in the electrode vessels, each of which contained one liter of buffer. The edge of each wick adjacent to the aluminum cooling plate was inserted into the soaked, slitted dialysis tubing so that a U-shaped covering over the wick was formed. Finally, an appropriate length of Saran Wrap was placed as a cover for the wick assembly. Without such covers, the wicks will dry to the point where they actually burn out (Peterson et al., 1966).

Prior to sample application, the 20 cm x 30 cm rectangle of paper was soaked in buffer, and blotted thoroughly between doubled sheets of Whatman 3MM filter paper. The sample was applied on a line 1.5 cm long perpendicular to the direction of electrophoresis. This sample line was located 2 cm from the chromatographic bottom of the

(7) Available from Calbiochem. P.O. Box 54282, Los Angeles, Calif. 90054

sheet, and 1.5 cm from the positive pole edge.

After sample application, the paper was transferred to the aluminum plate covered with a sheet of Mylar. The wicks were then adjusted to overlap the paper by at least .5 cm. The paper, with wicks overlain, was then covered by a second sheet of Mylar to prevent evaporation. A rectangle of Lucite (18 cm x 35 cm) was then placed on the Mylar over the paper, and weighted down with lead blocks. The protective cover of the Savant instrument(8) was closed, the constant voltage power supply turned on, and the voltage adjusted to 2800 volts. During a fourteen minute run, the current normally rose from an initial value of 150 milliamperes to a final value of 215 milliamperes. The temperature of the return cooling water seldom rose above 10°C from an initial value of 4°C. At the end of the electrophoresis, the paper was removed from the cell and allowed to air dry in a fume hood overnight.

The well-aired paper was sewn with cotton suture thread to form a cylinder and subjected to ascending chromatography at right angles to the direction of electrophoresis. The single phase solvent used for chromatography consisted of n-butanol, pyridine, glacial acetic acid, and deionized water in the ratio 15:10:3:12 (v/v) (Jegorov et al., 1965). The chromatography was performed in cylindrical jars 6" x 12", covered by a glass plate. When the solvent front had risen to within a centimeter or so of the top of the paper, the paper was removed from

(8) Model FP-18 high voltage electrophoresis cell complete with HV 3,000 constant voltage supply. Savant Instrument Inc. 221 Park Ave. Hicksville, N.Y.

the jar, the threads cut, and the paper hung in the hood to dry. Papers hung to dry in the cylindrical form tended to display distorted spots, probably due to a drying pattern. Such artifacts were seldom seen when the sheets were hung in a plane fashion after cutting the threads used to maintain the cylindrical form.

After drying overnight, the paper was again sewn into cylindrical form and subjected to a second chromatographic development in the same manner as before. Hoerman and Kamel (1967) and Jegorov et al., (1965) recommend multiple development for optimal resolution. In the case of the work presented here, double development enhanced the separation of those peptides possessing low R_f values.

The developed chromatogram was again dried in plane fashion. Those spots exhibiting fluorescence when exposed to ultraviolet light (2537 Å and 3660 Å)[9] were marked. Visualization of the peptide spots was accomplished by dipping the paper in a .02% (w/v) solution of ninhydrin in acetone and then allowing the color development to proceed at room temperature.

Once initial development had begun, the map was traced. After this tracing, the map was allowed to develop further overnight. The map tracing was then completed, the spots marked for identification and cut from the paper.

The peptides were eluted from the papers with 10% (v/v) formic acid, using filter paper between two microscope slides as wicks for a descending chromatographic elution. The peptides were eluted into

[9] An ultraviolet light source similar to the one used is available from Edmund Scientific Co. Barrington, N.J.

3" test tubes, permitting at least a milliliter of eluent to collect in the tube.

The labeled tubes, containing the eluted peptides were placed in a vacuum desiccator in the presence of separate containers of concentrated sulfuric acid and sodium hydroxide pellets. The desiccator was then evacuated to 5 mm Hg with an aspirator pump and left overnight.

The dried samples were removed from the desiccator and .15 ml of 5.6 N HCl was added to them. The resulting solution was then transferred into glass capillary tubes (diSpo pipettes).(10) The sample tubes were flame sealed at both ends and placed in a hot air oven at 110°C for 24 hours. When cooled, the tubes were opened at both ends and the samples washed out with a small volume of deionized water into 3" test tubes. The hydrolyzed peptide samples were evaporated to dryness in a vacuum desiccator containing concentrated sulfuric acid, sodium hydroxide pellets and Drierite in separate containers. The samples were stored when dry in a desiccator supplied with Drierite and sodium hydroxide pellets.

The above procedure results in the recovery of analytically acceptable quantities of material (.02 - .2 μM) from sample applications of four to five milligrams of digested protein. This is a recovery of the order of 10 to 70%, depending on the peptide.

Amino Acid Analysis

The amino acid analyses were performed using a system based

(10) Available from Scientific Products.

on the single column gradient elution ion exchange chromatography method of Piez and Morris (1960). The individual columns are essentially self-contained so that by use of sequential timers it is possible to perform analyses consecutively.

Once the system had been adjusted to satisfactory performance, standard runs were scheduled so that at least one standard was obtained for each column every two weeks. The standard mixture contained .5 μM/ml of each protein hydrolysate amino acid. The normal standard load was .1 μM or .2 ml of the standard solution. The standard also contained norleucine, an amino acid of non-biological origin which functioned as an internal standard since it was added to all samples. The norleucine concentration in the standard mixture and the sample buffer was 1.0 μM/ml. The elution program required an elapsed time of six and one-half hours for one complete analysis. Disodium ethylenedinitrilotetraacetate (EDTA) was incorporated in the buffers as recommended by Hamilton (1963) in order to overcome the accumulation of divalent ions which seriously impairs the performance of the columns of resin. Table 3 shows the degree of improvement seen after incorporation of this chelating agent into the buffers. In the course of standardizing the system, a number of problems were encountered; only two are of enough importance to be considered here.

One problem was the extreme variation of the recovery values for glutamic acid, aspartic acid and glycine. The first two amino acids were partially esterified by the methanol in the buffers, while the glycine variation was due to the fact that the glutamic acid gamma

Table 4

Effect of EDTA on recoveries from standard amino acid mixtures in terms of color equivalent (area) relative to norleucine, the internal standard.

	Without EDTA		With EDTA	
	Mean	C.V.	Mean	C.V.
Asp	.88	.05	.88	.06
Thr	.95	.04	.95	.02
Ser	1.02	.04	1.03	.02
Glu	.77	.11	.81	.11
Pro	.18	.17	.21	.07
Gly	1.14	.05	1.13	.05
Ala	.96	.04	.92	.03
Val	.87	.05	.85	.02
Cys	.50	.06	.52	.03
Met	1.05	.05	1.02	.02
Ile	.88	.03	.89	.05
Leu	.98	.04	.98	.02
Nor	1.00	-	1.00	-
Tyr	.93	.03	.93	.02
Phe	1.00	.04	1.00	.01
Lys	1.05	.06	1.03	.02
His	1.07	.06	1.04	.02
Arg	.98	.04	.96	.02
n=	29		10	

Note: These analyses were performed with buffers that contained methanol. The lack of improvement for glutamic and aspartic acids is attributed to ester formation described in a subsequent section.

methyl ester formed, was eluted in the position of glycine during chromatography.

The second problem was the desire to simplify the calculations of the data. The solution to this problem was found to be the use of peak height as the sole index of concentration in ion exchange amino acid chromatography. The methods by which this approach was established is included in this dissertation since such a method does represent a departure from commonly accepted procedures for computation of amino acid analysis data.

Degradation of Glutamic and Aspartic Acids

In the normal daily routine, samples were applied to three or four columns simultaneously. Unfortunately, therefore, except for the sample on the initial column which was immediately analyzed, the other samples remained absorbed to the resin at the tops of their columns for six to twenty-four hours during which time they were exposed to a pH of 2.75, a temperature of 60°C and a methanol containing (10% v/v) solvent system. The methanol was incorporated into buffers in order to improve the resolution of the two hydroxy-amino acids, threonine and serine (Thomson and Miles, 1964).

In the course of standardization of the total system, it was found that glutamic and aspartic acids varied beyond the acceptable $\pm$ 5% level. The values obtained, when compared or grouped on the basis of operational sequence, indicated no significant differences for the majority of the amino acids, but did indicate a progressive decrease in recovery of both glutamic and aspartic acids.

Figure 1 shows the results obtained for aspartic acid. Figure 2 shows a similar plot obtained for glutamic acid. The abscissa is in terms of order of analysis rather than time, since in the daily routine the sample for the first analysis of the day may remain on the column for up to six hours while the fourth analysis of the previous day is completed. The mean values, seen here, were used to correct each analysis on the basis of operational sequence, the values for the first sequence providing the base. These factors were included in a computer program which applied the correction on the basis of a sequence number given the analysis.

It was thought that esterification of the β carboxyl group of aspartic and the α carboxyl group of glutamic acid was the cause. The following work was performed to establish this and to search for a solution.

Two columns were loaded with 0.1 μM of glutamic acid,[11] the buffer above the resin bed contained methanol. Column #1 was developed immediately; that is, buffer was started through the column within thirty minutes after the initiation of sample application. Column #2 development was begun nineteen and a half hours later. Both columns produced two peaks, glutamic acid and a glycine elution time peak. The glycine elution time peak for column #1 amounted to 17.5% of the total area of the two peaks. For column #2 the glycine elution time peak amounted to 27.3% of the total area. A sample of

---

(11) The amino acids and amino acid esters used were obtained from nutritional Biochemicals Corp. The alcohols used were all Baker Analyzed Reagent grade. All buffers used contained disodium ethylenedinitrilotetraacetate.

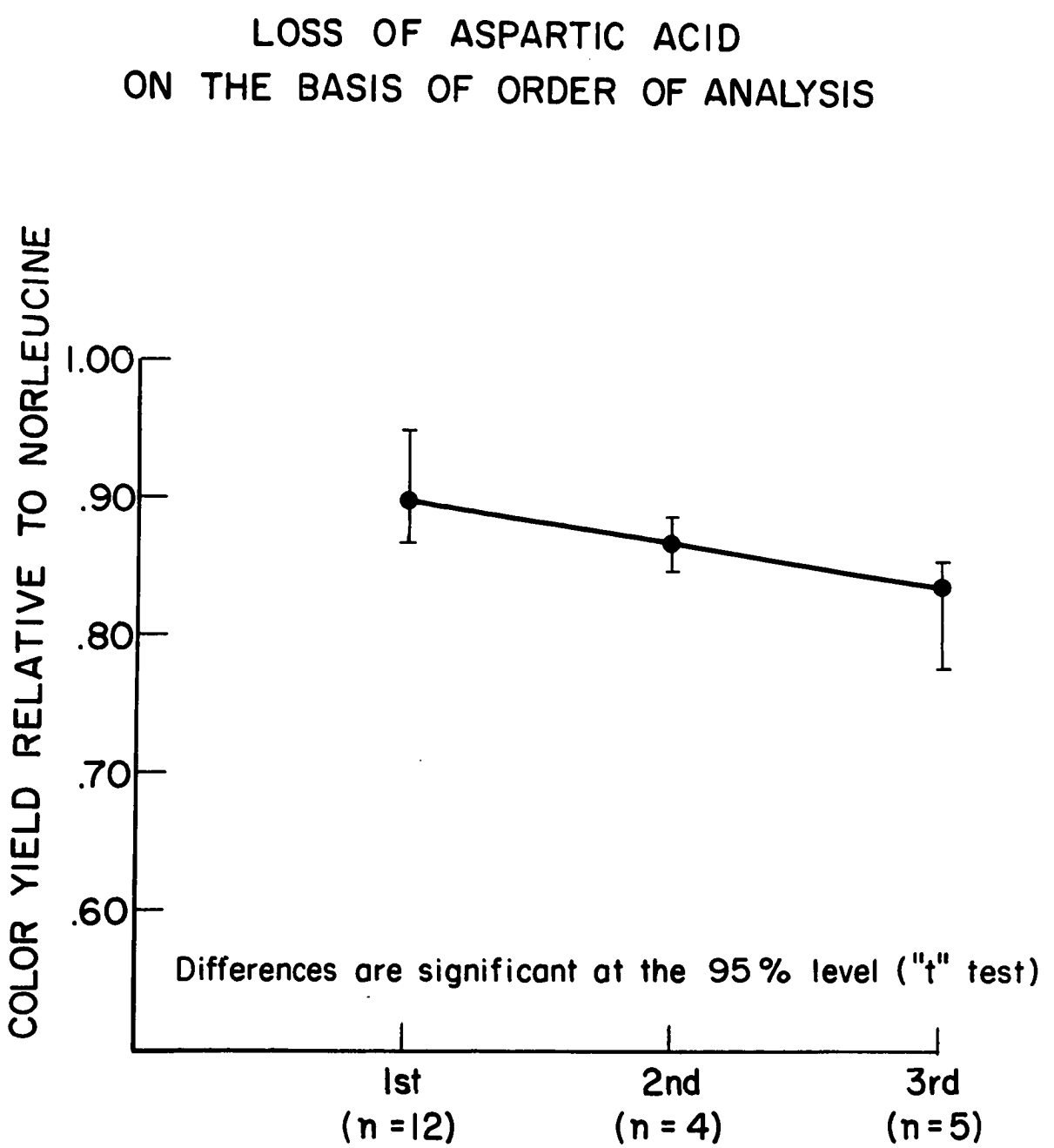

Figure 1. The progressive loss of aspartic acid on columns. The values plotted are the means and the range of values.

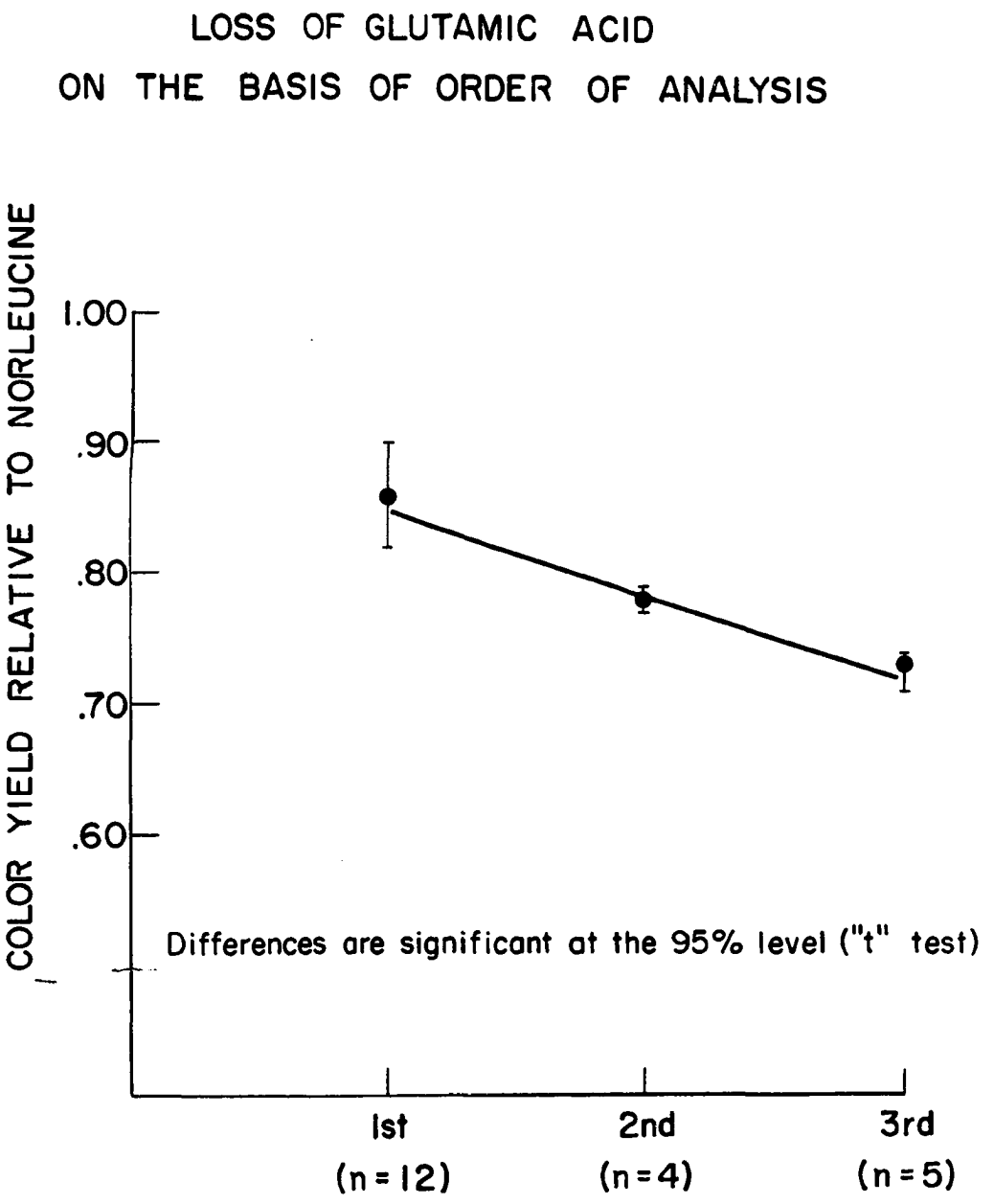

Figure 2. Graphical representation of the progressive loss of glutamic acid. The values plotted are the means and the range of values.

glutamic acid - gamma methyl ester was found to be eluted in the position of the glycine type peak. A sample of glutamic acid eluted without methanol in the buffers yielded a glutamic acid peak only.

The loss of aspartic acid was not as great as that of glutamic acid; however, it was found that the beta methyl derivative occurred in the serine area.

With glutamic acid, ethanol produced a second peak in the alanine region while n-propanol produced a second peak eluted after valine and close enough to cystine to cause a shoulder on that peak.

In both cases, a single peak was obtained, and this peak was eluted in the position predicted by elution of the same compound in the absence of any alcohol.

The threonine-serine resolution resulting from the use of tertiary butanol was at least equal to that provided by the use of methanol, with no apparent effect on the other portions of the chromatogram.

Gamma methyl glumatic acid was readily hydrolyzed to glutamic acid under the conditions employed. Difficulty was encountered in gaining recovery factors for it. The procedure finally adopted was to make up a solution in distilled water, apply the sample to the column and begin the run, all within fifteen minutes after making up the solution. Even with this rapid approach, a small glutamic acid peak amounting to 9% did occur. Chromatography attempted in a similar fashion in the absence of methanol resulted in a single peak in the glutamic acid region. It seems that the resin catalyzes this

conversion so that, in a very short time, the conversion is complete. Under such circumstances, only one peak would be observed and this would be attributed to the sample placed on the column. The above suggests that when methanol is present, the equilibrium for the gamma methylation is reached when 24 to 26% of the glutamic acid exists as the gamma methyl ester.

The situation with glutamic acid and aspartic acid appears clear cut. In the presence of methanol, some of the amino acid is converted to the gamma methyl or beta methyl ester. These compounds are then eluted in the areas of glycine and serine-asparagine respectively. This can be prevented by using tertiary butanol to replace methanol for the improvement of the resolution of threonine and serine.

Hamilton (1963) lists the position of gamma methyl glutamic acid in the region of glutamic acid. The above indicates that this ester behaves more like glycine. In the absence of methanol, this ester is recovered quantitatively as glutamic acid due to complete conversion of the ester to the parent acid. This probably explains the discrepancy in position, rather than invoking buffer schedule differences.

As a result of this work, the use of methanol in the buffers was discontinued and tertiary butanol was substituted. The sequence corrections for losses of glutamic acid and aspartic acid were no longer required.

Calculation of Amino Acid Composition

The graphical output from the amino acid analyzer is in the form of a series of absorbance peaks whose position and area correspond

to the kinds and amounts of amino acid present in the sample. In the original procedure for the automatic determination of amino acid composition, the area under the peak was computed by a triangulation procedure wherein the peak height (H), in terms of optical density units was multiplied by the width at half-height (W) in terms of dots, or time (Spackman et al. 1958). This is a time consuming method. In repetitive analyses, precision is the major concern, since the area obtained is subsequently converted by means of a factor into a mole value, the factor being derived from standards submitted to analysis at the same time as the sample under investigation.

Schroeder et al. (1966) reported the successful use of a nomogram for the calculation of amino acid data, but this approach requires frequent reconstruction of the chart based on current standardizations, due to minor fluctuations in pumping rates, buffer composition, and column performance. Mondino (1967) reported that peak height alone can be used successfully in the determination of amino acids eluted from ion-exchange columns where the ninhydrin reaction effluent is plotted by means of a linear response recorder, and the net height is estimated by means of a ruler. Mefferd et al. (1968) reported equal precision for seven different methods applied to 184 curves obtained by gas chromatography.

With this background, the precision obtained using peak height estimated manually was compared to that obtained using a Simpson's rule integration executed by an IBM 360 computer.

Table 5 lists the results from two sets of standards. The peak

Table 5

Comparison of values obtained by two methods of calculating amino acid analysis data

	(0.1μM) STANDARDS A *				(0.1μM) STANDARDS B **			
	Peak Height		Area		Peak Height		Area	
	Mean	C.V.	Mean	C.V.	Mean	C.V.	Mean	C.V.
Asp	0.85	0.03	0.82	0.02	0.98	0.05	0.79	0.05
Thr	0.92	0.02	0.90	0.01	1.10	0.03	0.94	0.04
Ser	0.99	0.03	0.97	0.02	1.12	0.02	1.04	0.04
Glu	0.82	0.05	0.94	0.01	0.78	0.04	0.83	0.05
Pro	0.19	0.04	0.20	0.04	0.19	0.04	0.20	0.05
Gly	1.29	0.03	1.20	0.02	1.30	0.07	1.38	0.06
Ala	0.85	0.05	0.85	0.01	0.77	0.04	0.88	0.04
Val	0.77	0.03	0.77	0.01	0.71	0.05	0.79	0.04
Cys	0.77	0.05	0.57	0.04	0.72	0.03	0.57	0.08
Met	1.09	0.02	1.11	0.01	1.04	0.03	1.14	0.08
Ile	0.83	0.02	0.77	0.02	0.83	0.02	0.79	0.04
Leu	1.10	0.01	1.07	0.02	1.10	0.02	1.11	0.02
Tyr	0.92	0.01	1.07	0.02	0.94	0.02	1.04	0.03
Phe	0.90	0.02	1.07	0.01	0.92	0.04	1.09	0.05
Lys	1.25	0.01	1.01	0.02	1.30	0.04	1.05	0.07
His	1.30	0.02	1.20	0.01	1.34	0.04	1.22	0.06
Arg	1.02	0.01	1.00	0.01	1.06	0.03	1.03	0.05
	N = 5				N = 7			

NOTES: *Standards A analyses were performed with tertiary butanol in the pH 2.75 buffer.

**The values for glutamic and aspartic acids were corrected for progressive loss since methanol was used.

height is expressed in optical density units as read from the printed chart. The base line values which were subtracted from the total optical density values were estimated on the basis of the base line before and after the peak. In the case of incompletely resolved peaks, the base line was interpolated. The area is the output from a computer program that performs the Simpson's rule integration for area based on punched tape digital information.

In both cases, the numbers obtained were normalized to the value for norleucine, which is routinely added to samples as an internal standard. The relative areas and relative peak heights for a number of analyses are tabulated. The data has not been grouped according to column, since experimental design is an attempt to randomize. The absolute values for color yield or recovery are different, but the precision is of the same order of magnitude.

Since the two methods appear equally precise for all practical purposes, the less laborious peak height procedure was adopted.

Losses of Tyrosine, Phenylalanine, and Histidine

The evaluation of the peptide mapping procedure utilized in this study was carried out using sperm whale myoglobin[12] as standard material. In the analyses of the tryptic peptides separated by this method, the recovery of tyrosine, phenylalanine, and histidine was of the order of 50% in some of the peptides containing these amino acids. It became apparant that where this low recovery was observed, the amino acid involved was amino terminal in the peptide. In the work

[12] Available from Gallard Schlesinger.

presented here, the loss seemed excessive and introduced some difficulty in conclusively identifying the peptides involved. In their work on harbor seal and porpoise myoglobin, Bradshaw et al. (1969) did not encounter similar low yields for these amino acids in the same peptides found here. For this reason, the method of mapping and acid hydrolysis (both different from the cited work) was suspected.

The low recovery of the amino terminal amino acid from peptides detected by reaction with ninhydrin on fingerprint maps such as those used here is not well reported or recorded in the literature. Bennett (1967) alludes to loss of amino terminal amino acids under these conditions. Genaux (1969) attributed this loss to the reaction with ninhydrin, which involves the amino terminal amino acid. Such reaction, an oxidative decarboxylation, necessarily leads to some loss of that particular amino acid. However, since losses of tyrosine, phenylalanine, and histidine are markedly greater than those seen for other amino acids occurring at the amino terminal position, they must be due to more than the reaction with ninhydrin.

These three amino acids all possess carbon-carbon double bonds and are therefore susceptible to oxidation. This suggests that the peptide mapping or the method of acid hydrolysis was the occasion for oxidation.

During the electrophoresis and chromatography, the peptides are spread over very large areas (up to 4 cm^2) and are exposed to an atmosphere that contains oxygen along with the volatile components of the developing solvent. Although electrophoresis lasts for only 14

minutes, chromatography extends over a 48 hour period at room temperature during which oxidation could well take place. Another possibility is oxidation during hydrolysis, since this step is carried out at 110°C for 24 hours, and during this study did not exclude oxygen by evacuation or displacement with nitrogen. As Blackburn (1968) points out this exclusion of oxygen is deemed essential to the quantitative recovery of those amino acids susceptible to such oxidation. The method of hydrolysis employed in this study dates back to earlier days and has been used by this writer successfully for a number of years.

In order to evaluate the impact of oxygen during hydrolysis, peptides T-5b (Phe-Lys), T-16 (Tyr-Lys), and T-6 (His-Leu-Lys) from standard sperm whale myoglobin were subjected to acid hydrolysis in tubes evacuated to 5 mm Hg with an aspirator pump. These samples were hydrolyzed in 5.6 N HCl at 110°C in an oil bath for 24 hours, in tubes constructed according to Schaffer et al. (1966). The globin preparations, the tryptic digestions, and the peptide mapping represented separate experiments.

Comparison of the data in Table 5 indicates that the low recoveries of tyrosine, phenylalanine, and histidine are not to be attributed to the presence of oxygen during hydrolysis. These results suggest that the loss occurs on the paper. Accordingly, it is most probable that the losses seen here occur during the chromatography (by oxidation) and visualization (by oxidative decarboxylation) on the paper.

Table 6

Comparison of recoveries of tyrosine, phenylalanine, and histidine in peptides hydrolyzed by two methods.

T - 5b		A	B
	Phe	.54	.54
	Lys	1.00	1.00
T - 16			
	Tyr	.44	.42
	Lys	1.00	1.00
T - 6			
	Leu	1.27	1.09
	His	.54	.58
	Lys	1.21	1.25

[A] The results of analyses of evacuated tubes

[B] That of non-evacuated

Mole Ratio Conversion of Data and Assignment of Peptides

The conversion of the data to a mole ratio basis is an application of the law of multiple proportions. In this case, instead of atomic composition, the integral combining proportions represent amino acid residues. This process of constructing peptides from compositional data is at present, unnamed and seldom, if ever, referred to in the literature. However, Schroeder (1968) does describe this process and provides a few examples.

The individual net peak heights were corrected for color yield or recovery by factors based on current standards. When the identity of the peptide is not known, it is not possible to normalize data for an analysis to some preconceived number of residues. Fortunately, there are usually a few amino acids present that occur as single residues in the peptide under consideration. By striking an average from these amino acids, a common divisor can often be found quickly. The values for those amino acids prone to loss are not utilized in the selection of this divisor. After this reduction of the data to a mole ratio form, the identity of the peptide may be determined. In the case of the myoglobin of sperm whale, a direct alignment due to identity was expected, and indeed, was found. In the case of the harbor seal myoglobin, it was expected that some identity would prevail and direct alignment would be possible, by homology, in some peptides at least. As will be seen, the peptides separated from Bering Sea harbor seal myoglobin provided data permitting direct alignment with the known sequence of Maine coast harbor seal myoglobin.

The principle of homology in its application here, takes the form "Identity of peptide composition indicates identity of peptide sequence". If a peptide isolated from a molecule whose sequence is unknown is identical in composition to a peptide from a corresponding protein whose sequence is known then the peptides are identical in sequence. Since actual sequence proof by way of chemical or enzymatic means was not to be performed, this assumption lies at the heart of the assignments that have been made. As more and more sequences are reported, it becomes increasingly obvious that sequencing by homology is a real source of error. Nolan and Margoliash (1968) note that the terminal tetrapeptide of the cytochromes c of duck, pigeon, and rattlesnake have the same composition, but the sequences are different. In the duck, it is Ala-Thr-Ala-Lys, in the pigeon, Thr-Ala-Ala-Lys, and in the rattlesnake, Lys-Thr-Ala-Ala. Nevertheless, in the work presented here, where the proteins of very closely related species are being compared, the assumption of homology seems quite reasonable, since only a few substitutions are expected for the whole molecule. It should be emphasized that it is an assumption and final proof of sequence must reside with actual sequence work.

Since there is no report in the literature for peptide maps of myoglobin using the conditions developed for this research, the identification of peptides presented a problem. Short peptides are easily identified, but larger ones are more difficult. The identification can be simplified by construction of a key in terms of specific amino acids, (Table 7). Such a key facilitates the proper use of time and

energy in the assignment of peptides so that troublesome analyses are quickly identified. Homology is basic to this key. Since both sperm whale myoglobin and harbor seal myoglobin were examined initially, the key includes the sequence for both, although for purposes of clarity, the harbor seal myoglobin peptides are emphasized.

The peptide designations generally follow those given for harbor seal myoglobin by Bradshaw et al. (1969b). The only change is the addition of certain sub-peptide designations, for example T-1c, T-1d, and T-1e. The total sequence of harbor seal myoglobin with designations is given in Figure 3.

Table 7. Peptide Identification Key

Lys/Arg	Pro	Met	Ile	His	Ala	Phe	Tyr		Peptide	(no. of residues)
Lys	+	+				+			T-13; 13a	(15)
	+	-	+	+					T-11	(6)
	+	-	+	-					T-11b	(4)
	+	-	-	-	+				T-10; a; b	(16-18)
	+	-	-	-	-				T-4	(8)
	-	+				-	-		T-7	(6) [S.W.]
	-	-	3	+	+	+	+		T-12	(16)
	-	-	+	-	+	-	-		T-15	(5, 6)
	-	-	-	+	+	-	-		T-8	(6) [S.W.]
	-	-	-	+	+	-	-		T-1	(16)
	-	-	-	+	+	-	-	F	T-1; 1e	(16); (5)
	-	-	-	+	-	-	-	Leu	T-6	(3)
	-	-	-	+	-	-	-	Ser	T-12c	(3)
	-	-	-	-	-		T-5b	(2)		
	-	-	-	-	-	+	-	Leu	T-3	(3)
	-	-	-	-	-	+	-	Asp	T-5a	(3)
	-	-	-	-	-	-	+		T-16	(2)
Arg		+							T-7	(6, 7) [H.S.]
		-	+						T-2b	(2)
			>1			+	+		T-12	(13)
			+			-	-		T-2	(15)
			-	+					T-14	(6)
			-	+		-	-		T-12c	(3)
			-	-		-	-		T-8	(5) [H.S.]
X X	X	X	3	+					T-12b	(9)
			-	+					T-1d	(4)
				+					T-9b	(6)
				+					T-2a	(13)
				-		+	+		T-12a	(4)
				-		-	-	F	T-1c	(7)

F = Peptide spot fluorescent under ultraviolet light

Table 8

One- and three-letter amino acid abbreviations.
(Dayhoff, 1969)

A/Ala = Alanine	M/Met = Methionine
C/Cys = Cysteine	N/Asn = Asparagine
D/Asp = Aspartic acid	P/Pro = Proline
E/Glu = Glutamic acid	Q/Gln = Glutamine
F/Phe = Phenylalanine	R/Arg = Arginine
G/Gly = Glycine	S/Ser = Serine
H/His = Histidine	T/Thr = Threonine
I/Ile = Isoleucine	V/Val = Valine
K/Lys = Lysine	W/Trp = Tryptophan
L/Leu = Leucine	Y/Tyr = Tyrosine
B/Asx = Aspartic acid or asparagine	Z/Glx = Glutamic acid or glutamine

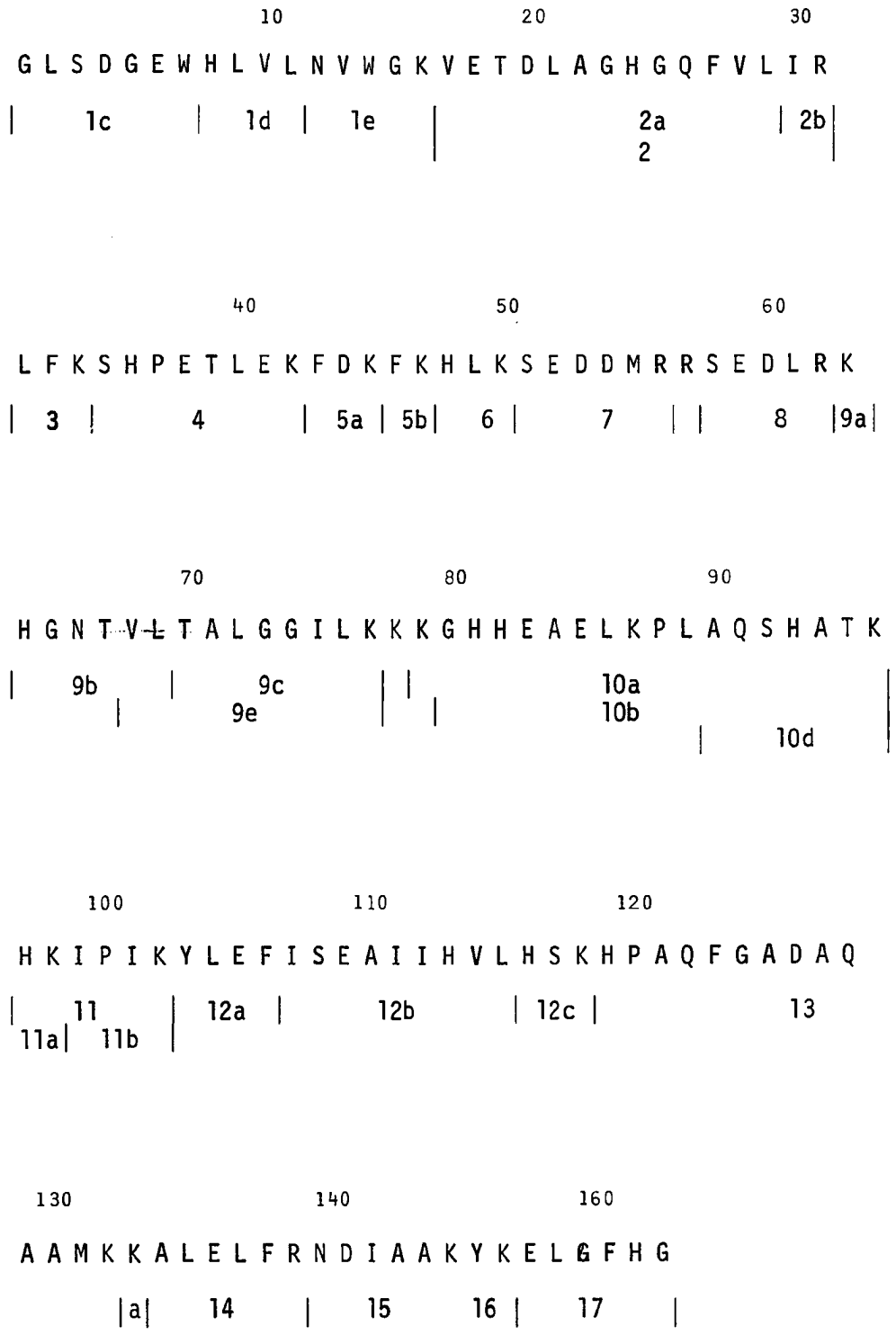

Figure 3. The amino acid sequence of harbor seal myoglobin as determined by Bradshaw et al. 1969b. The tryptic peptide designations are adapted from the same reference. The single letter code of Dayhoff is used and is given on the facing page.

RESULTS

The peptide maps usually contained 24 to 26 ninhydrin positive spots. In most cases, elution and hydrolysis of the peptides from these spots provided analytical data permitting identification of the peptides. The maps of phocid myoglobins generally permitted identification of all but two peptides, T-1e and T-12b. The same degree of success was not achieved for the myoglobins from walrus and fur seal.

This chapter is divided into two sections. The first section presents the peptide maps of the myoglobins of the six species of pinnipeds. The map of the peptides from sperm whale myoglobin is also included, since it provided the initial frame of reference. Once the analyses of the peptides from the myoglobin of the Bering Sea harbor seal had been completed, its map became the standard reference. The second section presents tabular data for the peptides eluted from the spots, together with comments on the analyses, and the sequence proposed for the peptide under consideration.

Peptide Maps of Tryptic Hydrolysates

As has been noted by many investigators, peptide maps are not strictly reproducible. Although the relative positions of the peptides do remain constant, identical samples run simultaneously or consecutively will produce maps with at least some minor differences. This problem is increased by the mapping method employed here. The small size of the paper adds to the general problem since it reduces the area available for separation of the peptide spots.

In the maps the dashes outline those spots that fluoresced when the

papers were exposed to ultraviolet light. The lines enclosing the ninhydrin positive areas were traced during early color development. The final spot was usually about two-fold larger, particularly in the direction of electrophoresis. Such elongation is due in part to overloading of the paper necessitated by the poor recovery of some of the peptides. In order to provide adequate recoveries for the identification of the amino acids, it was generally necessary to load the paper with digest from 4 to 5 milligrams of protein. Better resolution of some peptides would be possible if this loading could be reduced to 2 to 3 milligrams. Such a sample size is acceptable if the sensitivity of the amino acid analyzer can be increased. This is feasible either by range expansion of the recorder or the use of a longer flow cell. Both methods have been recommended (Hamilton, 1967) and are in present day use in many laboratories.

Other types of solid support media might provide higher recoveries, but to date, no generally accepted substitute for cellulose paper has been found for peptide work.

The small size of the paper, the high voltage gradients possible, and the overall convenience of the equipment involved more than made up for the difficulties enumerated above. The maps of the tryptic peptides of the pinniped myoglobins were very similar with many of the spots in the same position with only a few exceptions. The difference in the position of peptide T-1c of whale myoglobin relative to that from the pinniped myoglobins reflects the difference in composition of these two peptides. However, the position of T-4 was variable in the phocids, often leading to incomplete

resolution from T-15. The appearance of two T-7 spots was attributed to partial oxidation of the methionine present in the peptide. In the maps for sperm whale myoglobin, peptides T-7 and T-8 were found in a single spot, most probably as a single peptide, even though a peptide whose analysis resembled T-8a was found. Two spots, both of which contained only lysine are labeled 9a on all of the maps. The free amino acid, lysine, is liberated from positions #63, 78, and 133. The dipeptide, lysyl-lysine, could be liberated from positions #78 and 79 and having a lower charge to mass ratio should move more slowly in electrophoresis and possess a higher R_f in chromatography The arginine present at position #57 was not found as a part of T-8 nor was it found as the free amino acid. Peptide T-8 was found unresolved from T-13 in the maps of harbor seal, ribbon seal and bearded seal. Peptides T-8 and T-13 were not found in the walrus and fur seal maps. T-9b was sometimes poorly resolved from T-16 and T-9c was usually poorly resolved from T-14. In the maps for walrus and fur seal myoglobins, the spot for peptide T-9c was lacking, and a new spot corresponding to T-9e was found. In the maps from fur seal myoglobin a spot corresponding to peptide T-9b (with two amino acid replacements) was found close to peptide T-2 rather than T-14. The T-10 spots were often troublesome. The T-11b spot was usually rather faint and may have been overlooked in the early work on the phocid myoglobins, but since the parent peptide T-11 had been found for in all cases, the matter was not pursued. The peptide T-12c was found in the same position for all seven myoglobins, even though there was a substitution of glycine

for serine in the walrus and fur seal. Peptide T-13 was not recognized among the spots eluted from the maps of Weddell seal and bearded seal.

Peptide T-16 is constant in the phocids, but is found close to T-5b in fur seal and walrus. This difference in position was associated with the substitution of arginine for lysine in T-16.

Peptide T-17 occupies identical positions in the phocid maps reflecting identity of composition. Peptide T-17 of sperm whale, fur seal, and walrus myoglobins differ from that in the phocids both in map position and composition.

Once operating conditions had been established for the method, duplicate maps were usually prepared for each myoglobin under study. In general, this was sufficient to establish the composition of most of the peptides. Occasionally, it was necessary to prepare a third in order to clarify one or two points, for such reasons as sample loss or questionable analyses. Where no complications arose, two maps were sufficient to provide all the samples necessary. Since each map provided some 25 samples, it was inevitable that sample losses, analyzer difficulties, and other normal laboratory problems would make some repetition necessary.

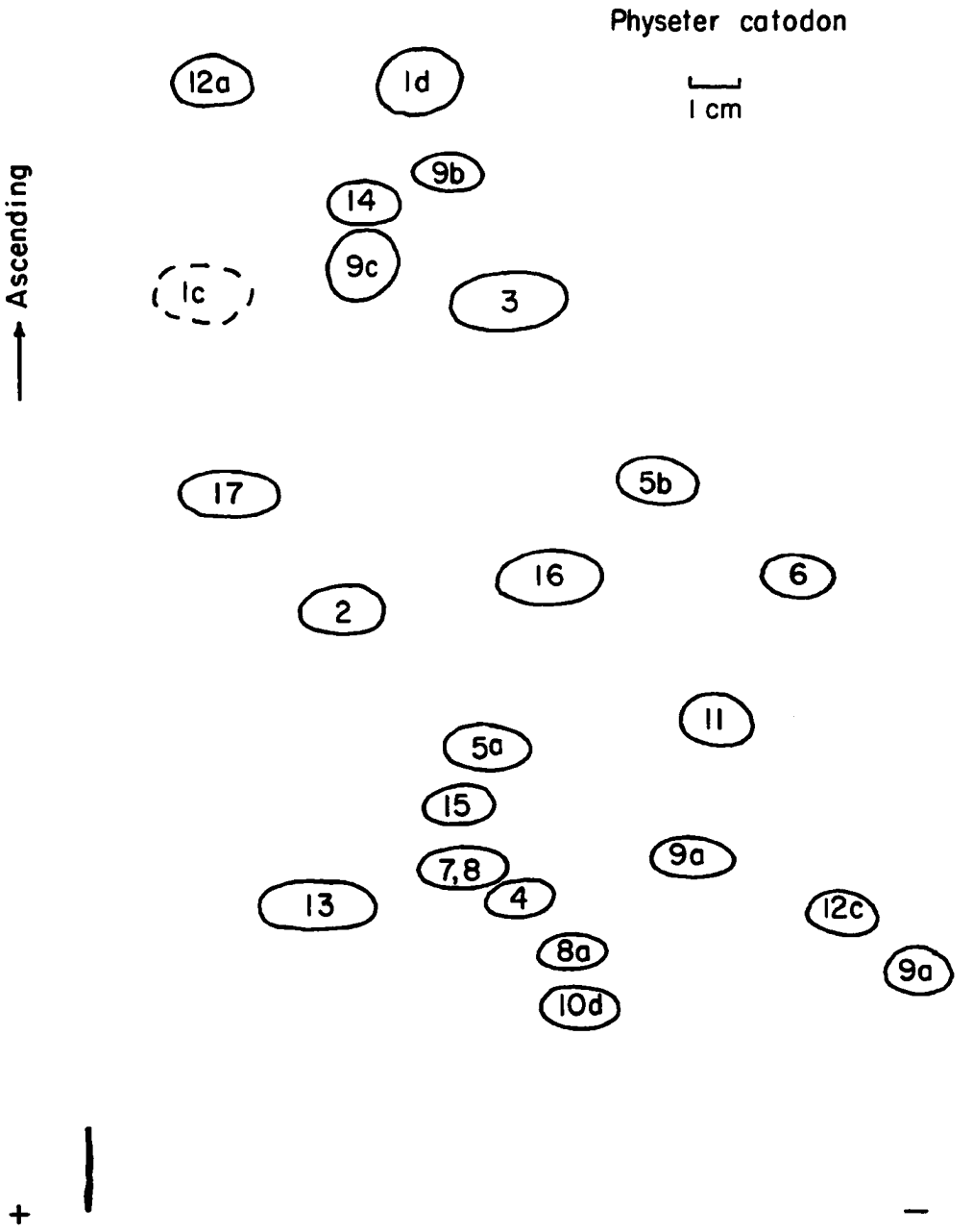

Figure 4. Tracing of the tryptic peptide map of sperm whale myoglobin.

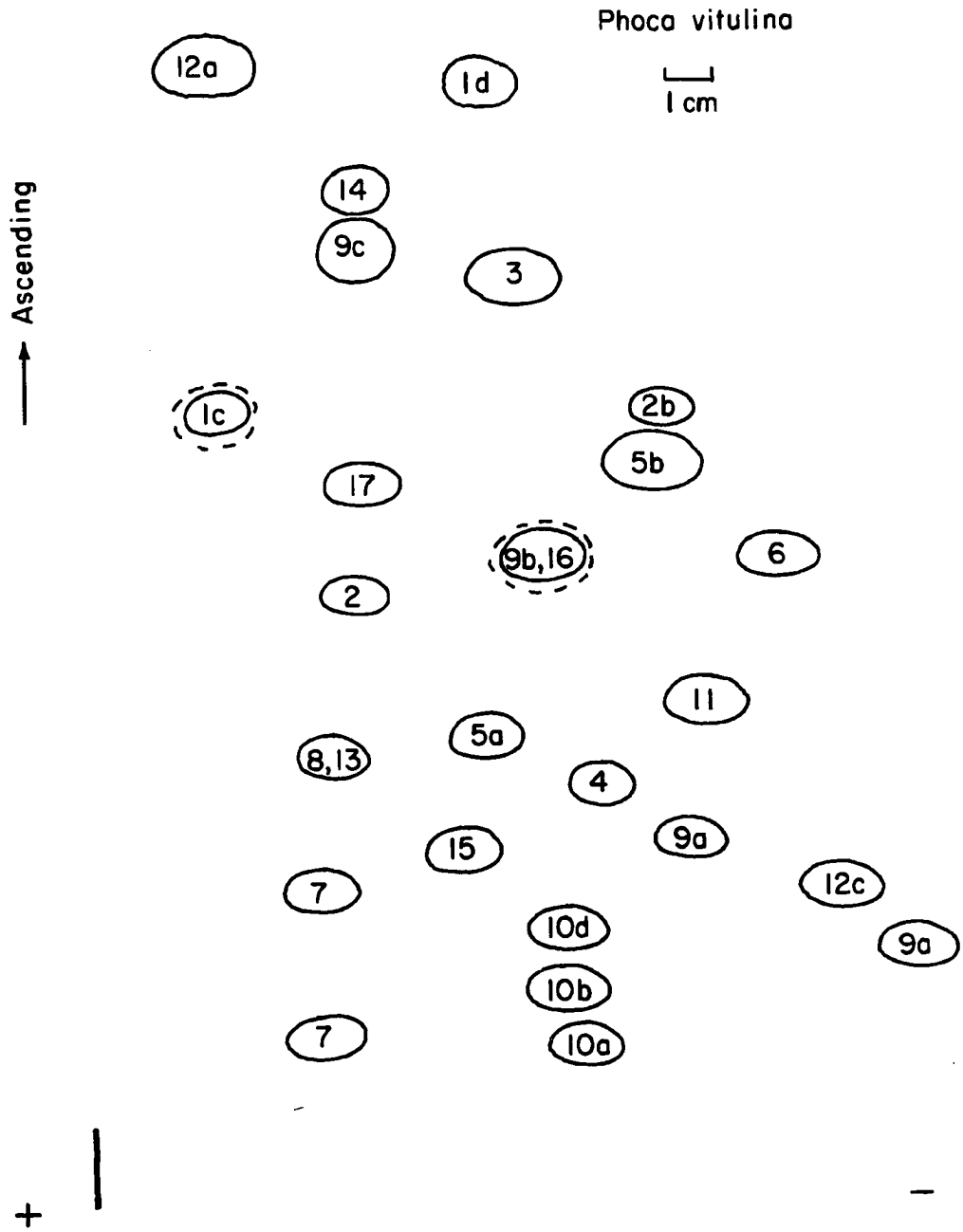

Figure 5. Tryptic peptide map of the myoglobin of the harbor seal.

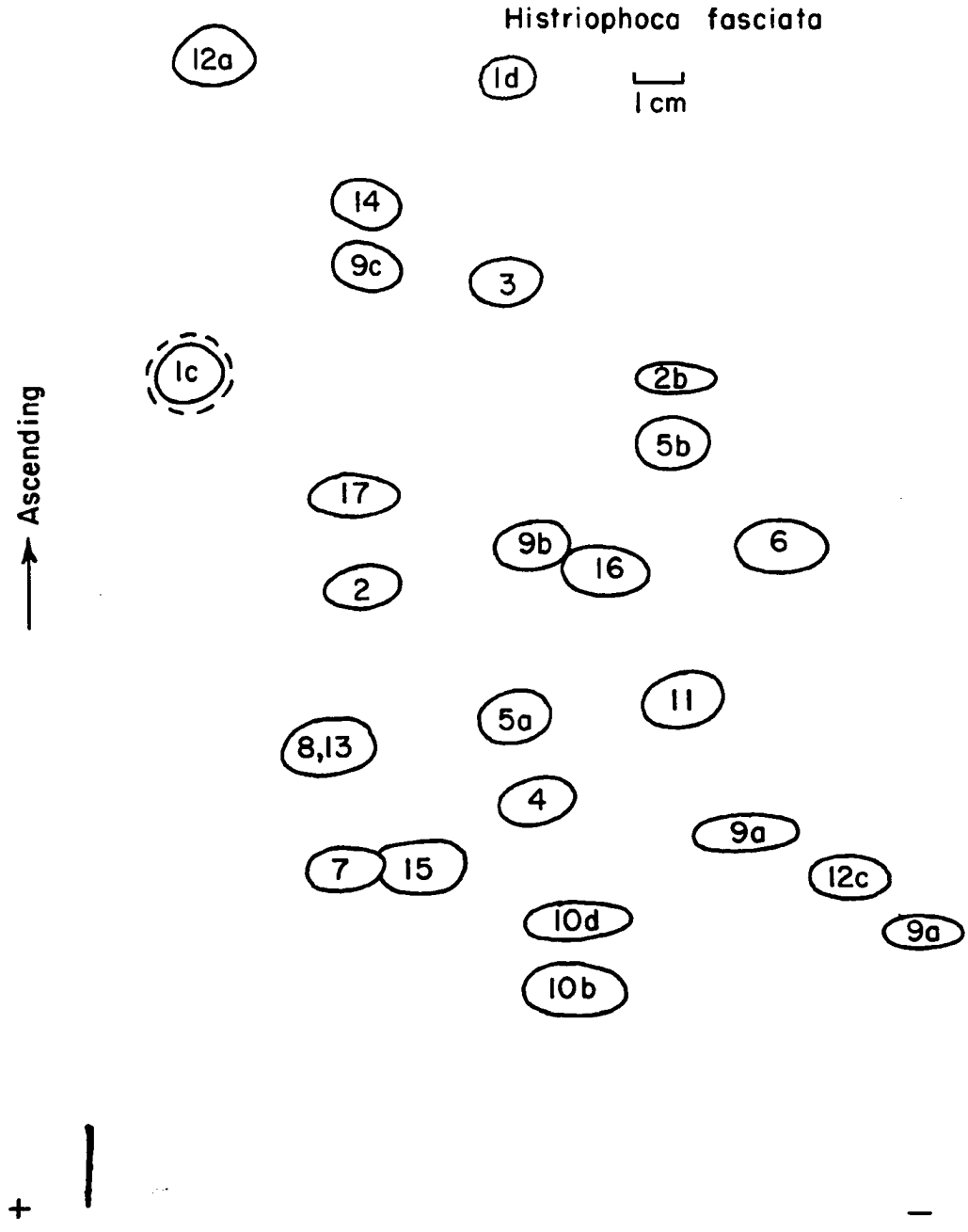

Figure 6. Tryptic peptide map of the myoglobin of the ribbon seal.

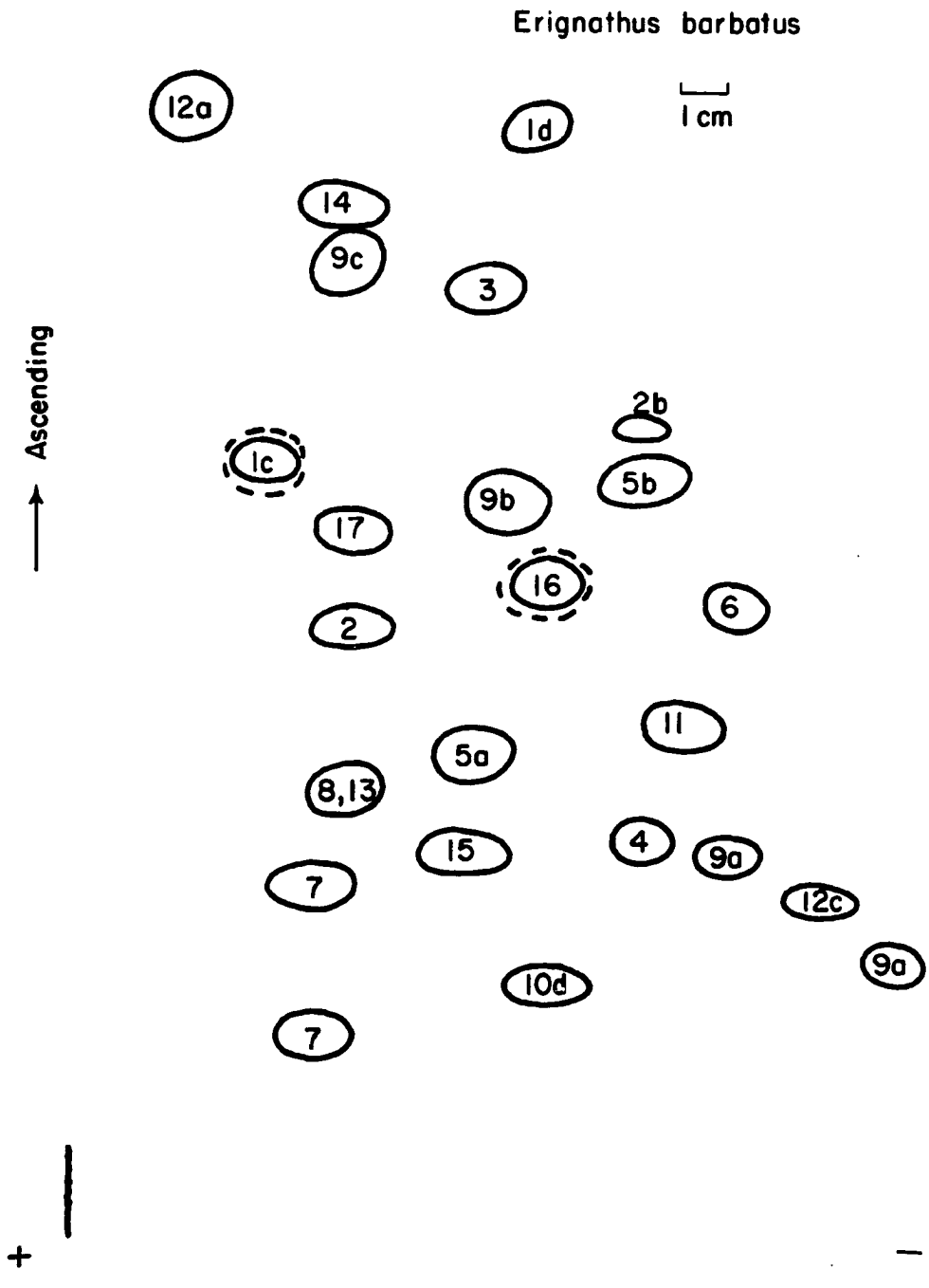

Figure 7. Tryptic peptide map of the myoglobin of the bearded seal.

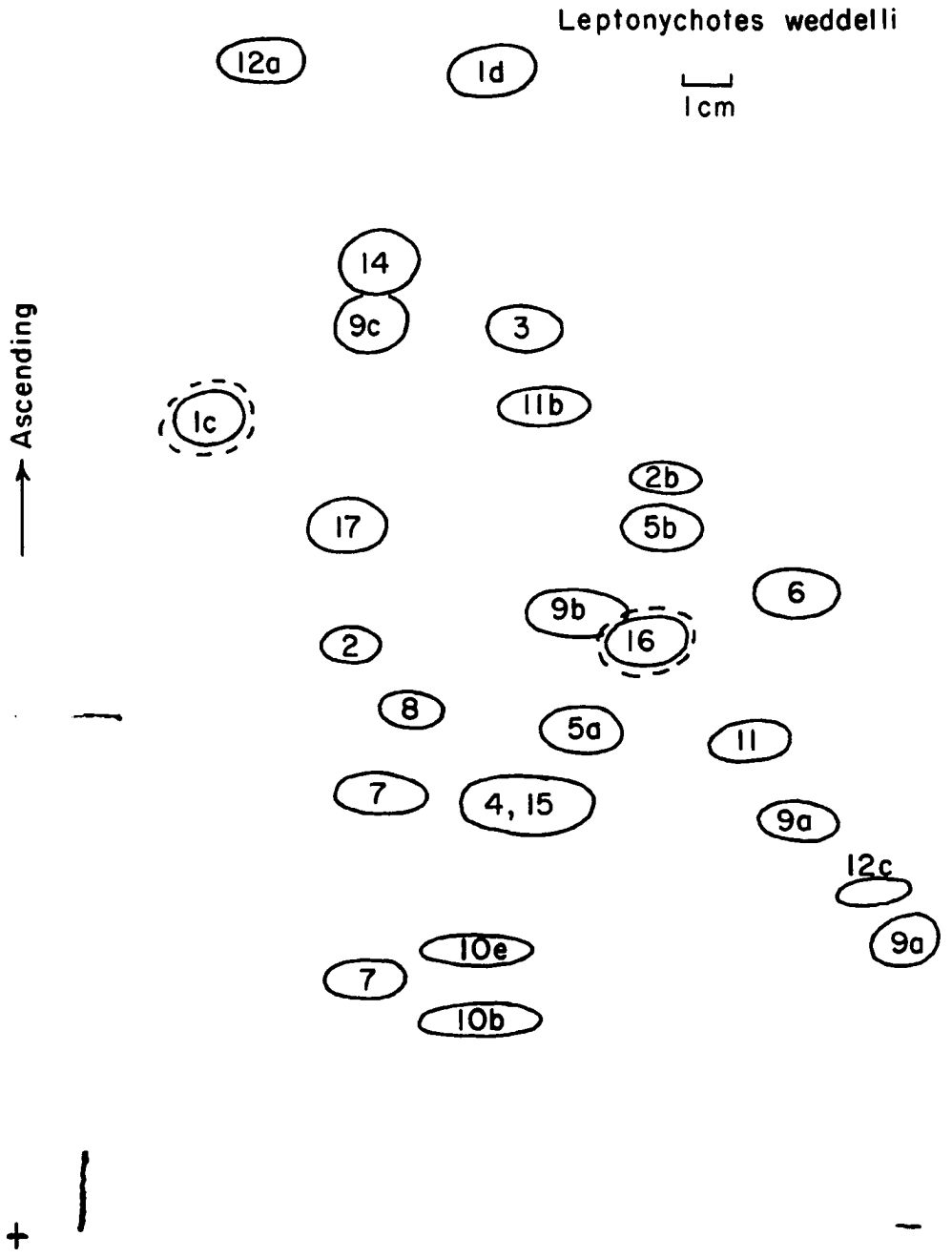

Figure 8. Tryptic peptide map of the myoglobin of the Weddell seal.

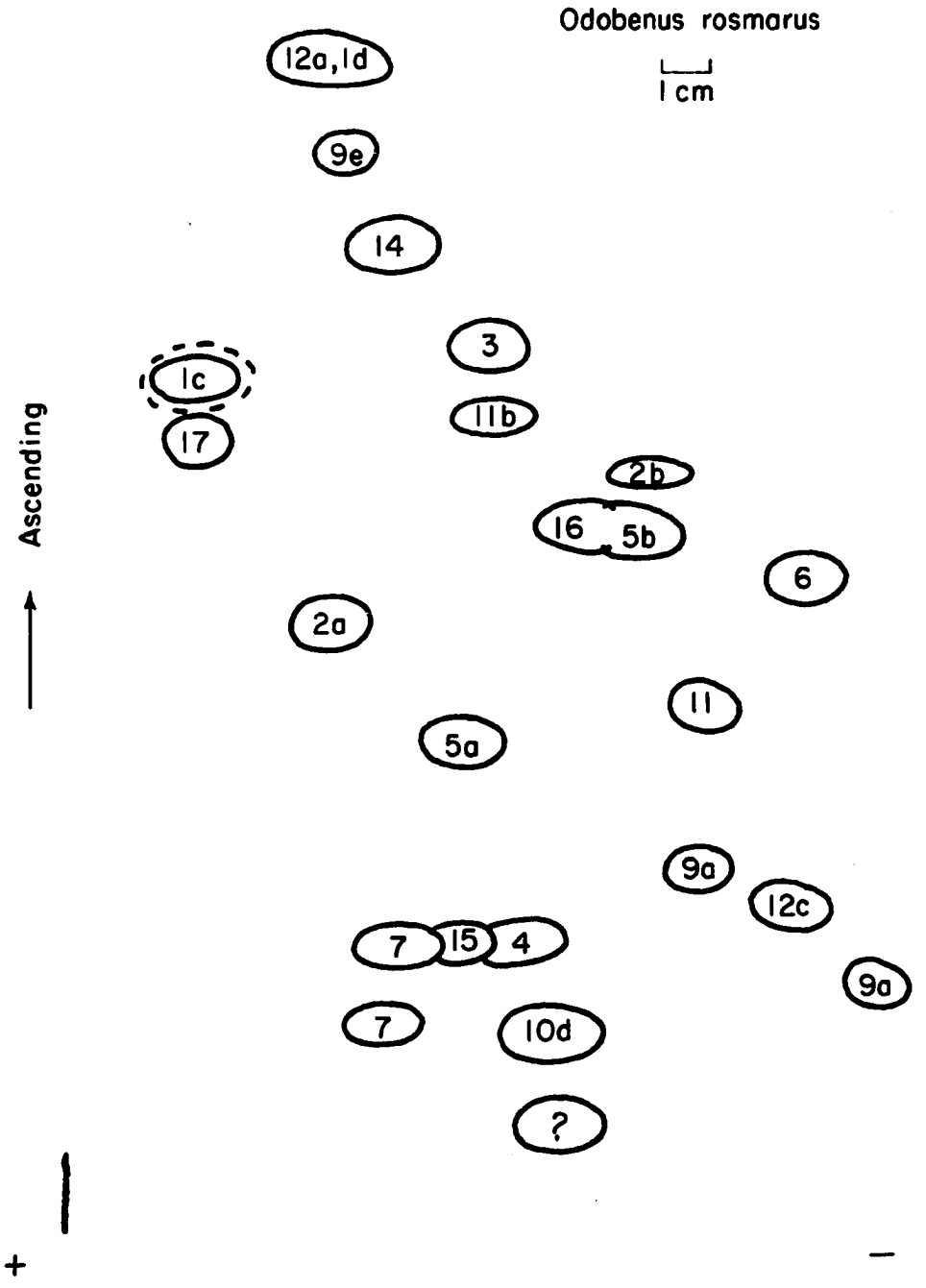

Figure 9. Tryptic peptide map of the myoglobin of the walrus.

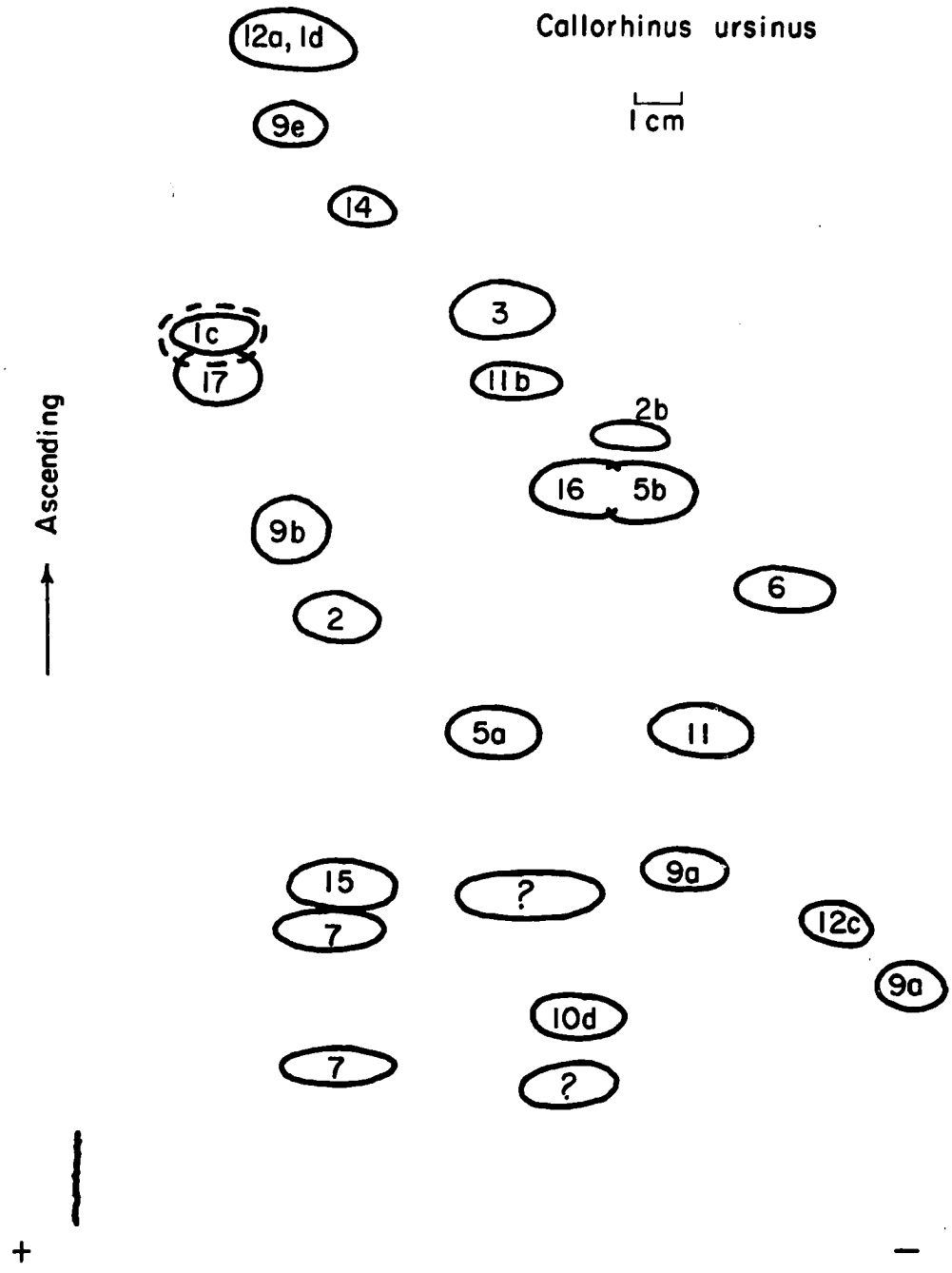

Figure 10. Tryptic peptide map of the myoglobin of the northern fur seal.

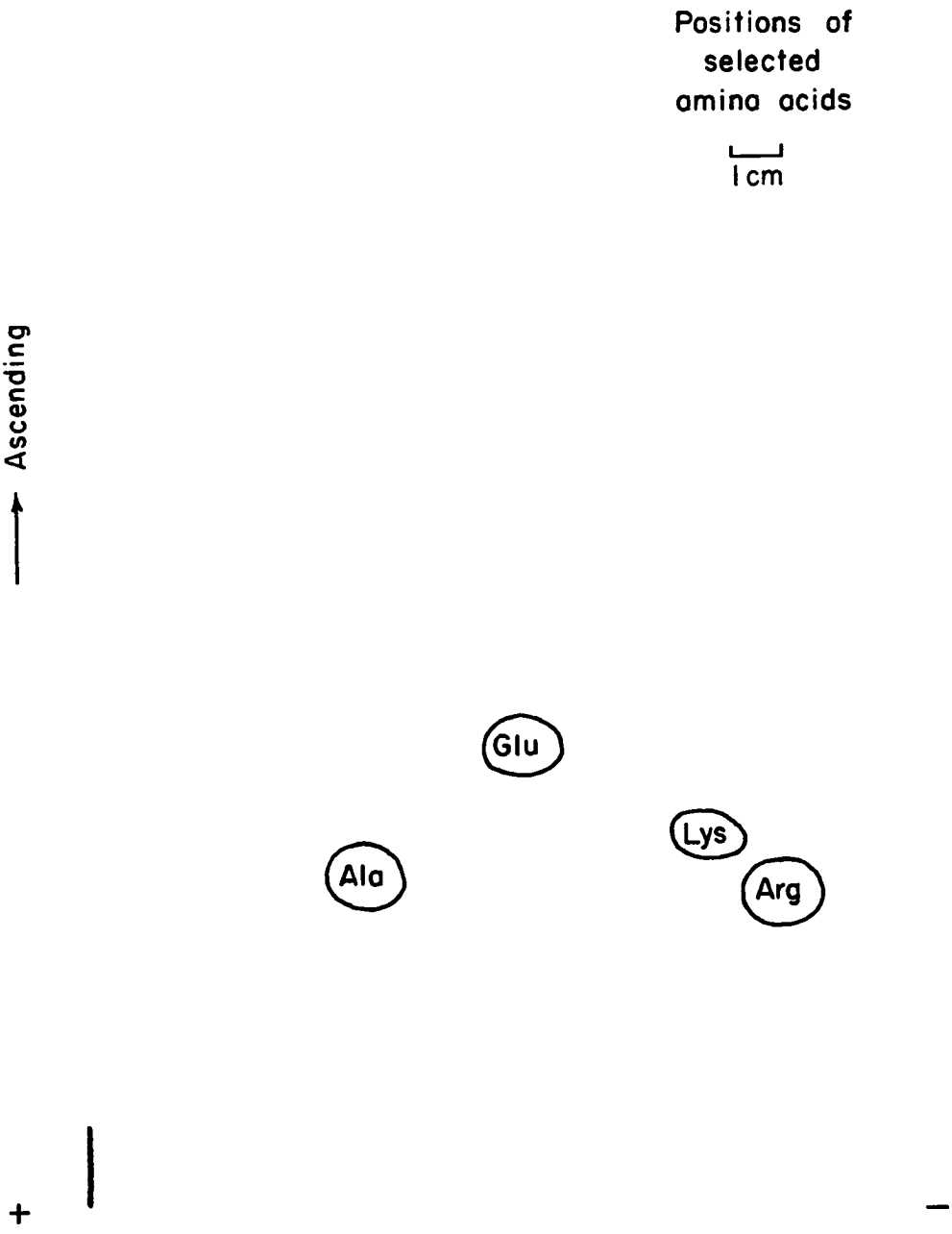

Figure 11.

The Analysis and Alignment of the Peptides

In the following comments, harbor seal, ribbon seal, bearded seal, and Weddell seal are sometimes grouped as "four phocids". The term "six pinnipeds" refers to the six species examined here. The alignments of the individual peptides are shown with the known sequence of harbor seal at the top and that of the sperm whale at the bottom.

Since glutamine and asparagine are hydrolyzed to glutamic acid and aspartic acid during acid hydrolysis of the peptides it is not possible on the basis of acid hydrolysis data to differentiate between the amide and the parent acid. Asx and Glx are used to indicate this uncertainty.

Proline, whose reaction with ninhydrin provides a color yield of only 20% of that of the other amino acids was sometimes poorly represented on an integral value basis. It was necessary to exercise judgment on occasion and state that this amino acid was present in an analysis, on the basis of appearance of the proline peak, relative to the balance of the other amino acids in the sample. Where proline's presence was in question it is indicated by a plus sign.

The peptides T-1, T-9 and T-12 are the intractable "core" peptides that usually resist the normal procedures for the separation of peptides. Edmundson and Hirs (1962c) treated this "core" material with chymotrypsin in order to establish the sequence for these three peptides. In the work presented here, it is seen that some of these chymotyptic peptides result from the tryptic digest employed for the fingerprinting. This was an unexpected advantage and permitted the identification more

of the molecule than was anticipated.

Incomplete separation of peptides is a recurring problem in fingerprinting work. In the case of column chromatography, there is usually sufficient material present to permit further manipulation to provide separation of the simpler mixtures sometimes found in chromatographic peaks. Where fingerprinting is performed in the manner utilized here there is not sufficient material available for further manipulation. Consequently, if unresolved peptides do occur the only attempt that can be made to separate them is a logical one. Initially, the maps of harbor seal myoglobins yielded one spot whose overall composition suggested a mixture of T-8 and T-13. Subsequently, with the pinniped maps, poor resolution was observed for other neighboring peptides on occasion. In some cases, a second or third map was prepared in order to achieve the resolution required. As many investigators have noted, the maps are not truly reproducible and as a consequence, the degree of resolution varies. As previously stated, it was necessary to overload the map slightly in order to obtain measurable yields of certain peptides. This tended to cause streaking in the electrophoretic direction, and this streaking, in turn, resulted in occasional poor resolution of some peptides. Peptides T-8 and T-13 were often unresolved as were T-4 and T-15.

The logical resolution of peptide mixtures from analytical data is open to question and, unless the matter was clear-cut the unresolved peptide analysis was listed as unclassified.

The peptide T-15 is rather unique in harbor seal myoglobin.

It is composed of two aspartic acids, one isoleucine, two alanine, and one lysine residue. This is readily separated from the mole ratio data presented in Table 9. Once the values for the peptide T-15 were recognized they were further improved to better reflect the composition of the peptide. Once the value for the peptide T-15 were recognized they were divided by .85. This improved the values so that they better reflected the composition of the peptide T-15. The remaining data resembled peptide T-4, and division by 1.40 provided values that were acceptable as peptide T-4.

This general procedure was used for the two sets of unresolved pairs encountered. A similar treatment of data for unresolved pairs is given by Schroeder (1968).

Peptide T-1c, Residues #1-7. This is the amino terminal portion of the molecule. Tryptophan was present since the spot from which this peptide was eluted fluoresced under ultraviolet light and lacks phenylalanine and tyrosine. The amino acid composition of this peptide was the same for all pinnipeds. The generally low recovery of glycine seen in the phocid peptide analyses is attributed to the fact that one glycine residue is amino terminal. By homology the myoglobins of all the pinnipeds examined are given the same sequence in positions #1-7 as that determined by Bradshaw et al. (1969b) for the harbor seal. This peptide differs from that of the sperm whale at positions #1 and #4.

Harbor seal	Gly-Leu-Ser-Asp-Gly-Glu-Trp
All pinnipeds	Gly-Leu-Ser-Asx-Gly-Glx-Trp
Sperm whale	Val-Leu-Ser-Glu-Gly-Glu-Trp

Table 9. Resolution of mixed peptide data from bearded seal peptide map.

AA	Mixed Data	T-15 Raw	T-15 Improved	T-4 Raw	T-4 Improved
Asp	1.66	1.66	1.96 ≃ 2		
Thr	1.14			1.14	.81 ≃ 1
Ser	1.04			1.04	.74 ≃ 1
Glu	3.15			3.15	2.25 ≃ 2
Pro	1.23			1.23	.88 ≃ 1
Gly	.43			.43	.31
Ala	1.96	1.96	2.31 ≃ 2		
Met	.12			.12	
Ile	.70	.70	.83 ≃ 1		
Leu	1.60			1.60	1.14 ≃ 1
Phe	.18			.18	
Lys	2.20	.70	.83 ≃ 1	1.50	1.07 ≃ 1
His	1.47			1.47	1.05 ≃ 1
Arg	.47			.47	.33

Peptide T-1d, Residues #8-11. In two of the phocids, this peptide provided reasonable values if the low recovery for histidine is accepted as due to its amino terminal position. In the harbor seal and the ribbon seal, the analytical values provided only qualitative evidence. In the bearded seal and the Weddell seal, a spot in the same area of the peptide map appeared to correspond to peptide T-1d in the harbor seal sequence. The recoveries were low (of the order of .01 μM) and hence, the analyses were not definitive. Repeated attempts did not materially improve the recovery.

A spot similarly located in the peptide maps of the myoglobins of the fur seal and the walrus provided evidence that the sequence of myoglobin in these two species was directly alignable with that of the sperm whale. In the sperm whale, fur seal, and walrus maps, there was serious contamination. Since the contaminants are similar, it is provisionally concluded that this peptide from the fur seal and walrus possess the same composition and, therefore, sequence as that of the sperm whale.

Harbor seal	His-Leu-Val-Leu
Four phocids	His-Leu-Val-Leu
Walrus and fur seal	Glx-Leu-Val-Leu
Sperm whale	Glu-Leu-Val-Leu

Peptide T-1e, Residues #12-16. This peptide was not found in any of the maps.

Peptide T-2, Residues #17-31. This is the second largest peptide recovered. The analyses of the phocid peptides showed identical

composition.

The T-2 spots from the maps of walrus and fur seal myoglobins were troublesome. Repeated attempts to obtain clear-cut analyses were only partially successful. One walrus myoglobin map did provide an analysis that corresponded well to peptide T-2a. This was the only time this peptide was obtained. The T-2 spot from fur seal myoglobin did not provide a clear cut analysis, and the best data is presented in the table to indicate the nature of the problem. This peptide is probably contaminated with a glycyllysine peptide arising from positions #15 and #16 since such a peptide was found in tryptic digests of harbor seal myoglobin by Bradshaw et al. (1969b). The final two residues of T-2 (T-2b) were found, showing that the lysine present is not at position #31. It is not possible to consider this lysine as a contaminant as the free amino acid - because lysine is found in an entirely different position.

Harbor seal	Val-Glu-Thr-Asp-Leu-Ala-Gly-His-Gly-Gln-Glu-Val-Leu-Ile-Arg
Phocids	Val-Glx-Thr-Asx-Leu-Ala-Gly-His-Gly-Glx-Glx-Val-Leu-Ile-Arg
Walrus	Val-Glx-Ala-Asx-Leu-Ala-Gly-His-Gly-Glx-Glx-Val-Leu
Sperm whale	Val-Glu-Ala-Asp-Val-Ala-Gly-His-Gly-Gln-Asp-Ile-Leu-Ile-Arg

Peptide T-2b, Residues #30-31. This fragment from the carboxyl terminal portion of peptide T-2 was usually found without difficulty

Table 10. Amino acid residue composition of peptide T- 1c (Residues #1-7)

	Phocidae				Odobenidae	Otariidae	Cetacea
AA	*Phoca* exp. ref.	*Histrio-phoca* exp.	*Erig-nathus* exp.	*Leptony-chotes* exp.	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp. ref.
Asp	1.03 ≃ 1	.94	1.01	.82	.71 ≃ 1	.81 ≃ 1	
Thr							
Ser	1.00 ≃ 1	.72	1.07	.83	.64 ≃ 1	.83 ≃ 1	.89 ≃ 1
Glu	.88 ≃ 1	1.17	1.19	.83	1.35 ≃ 1	1.26 ≃ 1	1.82 ≃ 2
Pro							
Gly	1.73 ≃ 2	1.68	1.83	1.83	2.14 ≃ 2	1.69 ≃ 2	1.27 ≃ 1
Ala							
Val							.89 ≃ 1
Met							
Ile							
Leu	1.00 ≃ 1	.88	1.06	1.04	1.14 ≃ 1	1.00 ≃ 1	1.09 ≃ 1
Tyr							
Phe							
Lys							
His							
Arg							
Trp	+ ≃ 1	+	+	+	+ ≃ 1	+ ≃ 1	+ ≃ 1

Table 11. Amino acid residue composition of peptide T-1d (Residues #8-11)

	Phocidae				Odobenidae	Otariidae	Cetacea	
	Phoca		*Histriophoca*	*Erignathus*	*Leptonychotes*	*Odobenus*	*Callorhinus*	*Physeter*	
AA	exp.	ref.	exp.	exp.	exp.	exp.	exp.	exp.	ref.
Asp									
Thr							.40		
Ser									
Glu						.88	≃ 1	.90 ≃ 1	.88 ≃ 1
Pro									
Gly							.39	.44	
Ala							.45	.62	
Val	+	≃ 1	+	1.04	1.07	.98	≃ 1	.93 ≃ 1	.86 ≃ 1
Met									
Ile									
Leu	+	≃ 2	+	2.05	2.08	2.00	≃ 2	1.95 ≃ 2	2.00 ≃ 2
Tyr									
Phe							.53	.72	
Lys									
His	+	1	+	.38	.70				
Arg									
Trp									

Table 12. Amino acid residue composition of peptide T-2 (Residues #17-31)

	Phocidae					Odobenidae*		Otariidae	Cetacea	
	Phoca			*Histrio-phoca*	*Erig-nathus*	*Leptony-chotes*	*Odobenus*		*Callorhinus*	*Physeter*	
AA	exp.	≃	ref.	exp.	exp.	exp.	exp.	≃	exp.	exp.	≃ ref.
Asp	1.13	≃	1	1.01	1.24	1.22	.85	≃ 1	1.22	1.80	≃ 2
Thr	.96	≃	1	.79	1.07	1.03					
Ser									.72		
Glu	2.68	≃	3	2.71	2192	2.80	3.30	≃ 3	2.89	2.17	≃ 2
Pro											
Gly	2.40	≃	2	2.08	2.10	2.12	1.97	≃ 2	1.83	1.75	≃ 2
Ala	.98	≃	1	.99	1.17	1.30	1.92	≃ 2	1.78	2.17	≃ 2
Val	1.71	≃	2	1.70	1.60	1.60	1.79	≃ 2	1.88	2.12	≃ 2
Met											
Ile	.99	≃	1	.72	1.01	1.00			.79	2.10	≃ 2
Leu	2.07	≃	2	1.73	2.26	2.04	2.10	≃ 2	2.20	1.20	≃ 1
Tyr											
Phe											
Lys	.43			.44	.35				1.16		
His	.91	≃	1	.94	.93	1.00	.88	≃ 1	.79	.90	≃ 1
Arg	.90	≃	1	.99	.93	.97			.69	.98	≃ 1
Trp											

* This is peptide T-2a (Residues #17-29)

Table 13. Amino acid residue composition of peptide T-2b (Residues #30-31)

	Phocidae				Odobenidae	Otariidae	Cetacea	
	Phoca		*Histrio-phoca*	*Erig-nathus*	*Leptony-chotes*	*Odobenus*	*Callorhinus*	*Physeter*	
AA	exp.	ref.	exp.	exp.	exp.	exp.	exp.	exp.	ref.
Asp									
Thr									
Ser									
Glu									
Pro									
Gly							.33		
Ala									
Val									
Met									
Ile	.73	≃ 1	.91	.95	.89	.92	≃ 1	.86 ≃ 1	
Leu									
Tyr									
Phe									
Lys						.41		.46	
His									
Arg	1.00	≃ 1	1.16	1.05	1.10	1.04	≃ 1	1.00 ≃ 1	
Trp									

in the pinniped maps, but was not found in the sperm whale myoglobin maps. Its recovery on the fur seal peptide map indicates that the carboxyl terminal amino acid of that peptide is arginine rather than lysine but the contamination by glycine and lysine is unexplained.

Harbor seal	Ile-Arg
Phocids	Ile-Arg
Walrus Fur seal	Ile-Arg

Peptide T-3, Residues #32-34. This peptide is identical in all species examined.

Harbor seal	Leu-Phe-Lys
All pinnipeds	Leu-Phe-Lys
Sperm whale	Leu-Phe-Lys

Peptide T-4, Residues #35-42. This peptide was mixed with T-15 in the phocid maps examined.

The alanine contamination of sperm whale T-4 is unexplained. It could be due to free alanine from position #90 released by chymotryptic cleavage. The free amino acid alanine does occur in this general area as is seen from the tracing of the map of selected amino acids.

In the fur seal and walrus maps, the comparable peptide spot yielded an analysis providing for recognition of T-15 but the remaining data did not resemble a typical T-4 peptide. The data remaining from the walrus analysis suggests that both threonine and serine have been replaced by glycine and alanine. In horse and bovine myoglobins, position #35 is occupied by glycine (Dautrevaux et al.,

1969). The low recovery of histidine is unexplained and the sequence proposed is questionable. This sequence is not included in the one proposed for walrus and fur seal myoglobins since the fur seal data was less conclusive and is not included at all.

Harbor seal	Ser-His-Pro-Glu-Thr-Leu-Glu-Lys
Phocids	Ser-His-Pro-Glx-Thr-Leu-Glx-Lys
Walrus	Gly-His-Pro-Glx-Ala Leu-Glx-Lys
Sperm whale	Ser-His-Pro-Glu-Thr-Leu-Glx-Lys

Peptide T-5a, Residues #43-45. The recovery of the amino terminal phenylalanine is poor. While this peptide was seriously contaminated in all but the sperm whale, it is evident that the substitution of lysine for arginine seen in the harbor seal persists through those pinnipeds examined.

Harbor seal	Phe-Asp-Lys
All pinnipeds	Phe-Asx-Lys
Sperm whale	Phe-Asp-Arg

Peptide T-5b, Residues #46-47. This peptide from myoglobins of fur seal and walrus was occasionally poorly resolved from the corresponding T-16 peptide tyrosine-arginine. Nevertheless, the composition of this peptide is identical for the species examined.

Harbor seal	Phe-Lys
All pinnipeds	Phe-Lys
Sperm whale	Phe-Lys

Peptide T-6, Residues #48-50. The consistent values obtained for this peptide indicate identity for this sequence in all species examined.

Table 14 Amino acid residue composition of peptide T-3 (Residues #32-34)

	Phocidae				Odobenidae	Otariidae	Cetacea	
AA	*Phoca* exp.	ref.	*Histrio-phoca* exp.	*Erig-nathus* exp.	*Leptony-chotes* exp.	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp.	ref.
Asp									
Thr									
Ser									
Glu									
Pro									
Gly									
Ala									
Val									
Met									
Ile									
Leu	.79	≃ 1	.86	1.00	.85	.88 ≃ 1	.89 ≃ 1	.77	≃ 1
Tyr									
Phe	1.16	≃ 1	.97	.86	1.00	1.02 ≃ 1	1.01 ≃ 1	1.00	≃ 1
Lys	1.05	≃ 1	1.28	1.13	1.05	1.12 ≃ 1	1.03 ≃ 1	1.25	≃ 1
His									
Arg									
Trp									

Table 15 Amino acid residue composition of peptide T-4 (Residues #35-42)

	Phocidae				Odobenidae	Otariidae	Cetacea	
AA	*Phoca* exp.* ref.	*Histrio-phoca* exp.	*Erig-nathus* exp*	*Leptony-chotes* exp.*	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp.	ref.
Asp								
Thr	.73 ≃ 1	.72	.81	.79			.78	≃ 1
Ser	.77 ≃ 1	.87	.74	.79			.61	≃ 1
Glu	2.12 ≃ 2	2.11	2.25	2.30	2.18		1.91	≃ 2
Pro	.83 ≃ 1	.68	.88	.75	+		.63	≃ 1
Gly				.61				
Ala				.69			.63	
Val								
Met								
Ile								
Leu	1.19 ≃ 1	1.14	1.14	1.30	1.01		1.04	≃ 1
Tyr								
Phe								
Lys	1.16 ≃ 1	1.17	1.07	1.00	1.06		1.24	≃ 1
His	.86 ≃ 1	.83	1.05	.71	.42		1.00	≃ 1
Arg								
Trp								

*Data derived from unresolved pair.

Table 16 Amino acid residue composition of peptide T-5a (Residues #43-45)

	Phocidae				Odobenidae	Otariidae	Cetacea	
AA	*Phoca* exp.	ref.	*Histrio-phoca* exp.	*Erig-nathus* exp.	*Leptony-chotes* exp.	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp.	ref.
Asp Thr Ser	.99	≃ 1	.99	.79	.99	.91 ≃ 1	1.03 ≃ 1	1.14	≃ 1
Glu Pro Gly			.54	.39	.53	.34	.46 .34		
Ala Val Met	.45		.39			.41	.48		
Ile Leu Tyr	.51		.46		.53	.46	.48		
Phe	.58	≃ 1	.43	.67	.72	.60 ≃ 1	.59 ≃ 1	.68	≃ 1
Lys	1.10	≃ 1	1.02	1.37	1.09	1.00 ≃ 1	1.09 ≃ 1		
His	.34		.34	.38	.34		.36		
Arg Trp								1.19	≃ 1

Table 17 Amino acid residue composition of peptide T- 5b (Residues #46-47)

	Phocidae				Odobenidae	Otariidae	Cetacea	
AA	*Phoca* exp.	ref.	*Histrio-phoca* exp.	*Erig-nathus* exp.	*Leptony-chotes* exp.	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp.	ref.
Asp									
Thr									
Ser									
Glu									
Pro									
Gly									
Ala									
Val									
Met									
Ile									
Leu									
Tyr									
Phe	.72	≃ 1	.69	.61	.78	.78 ≃ 1	.56 ≃ 1	.72	≃ 1
Lys	1.28	≃ 1	1.28	1.00	1.25	1.08 ≃ 1	1.09 ≃ 1	1.32	≃ 1
His									
Arg									
Trp									

Table 18 Amino acid residue composition of peptide T-6 (Residues #48-50)

	Phocidae				Odobenidae	Otariidae	Cetacea	
	Phoca		*Histrio-phoca*	*Erig-nathus*	*Leptony-chotes*	*Odobenus*	*Callorhinus*	*Physeter*	
AA	exp.	ref.	exp.	exp.	exp.	exp.	exp.	exp.	ref.
Asp									
Thr									
Ser									
Glu									
Pro									
Gly									
Ala									
Val									
Met									
Ile									
Leu	1.00	≃ 1	1.19	1.02	1.10	.98	≃ 1	.97	≃ 1	1.09	≃ 1
Tyr									
Phe									
Lys	1.10	≃ 1	1.19	1.13	1.23	1.16	≃ 1	1.15	≃ 1	1.25	≃ 1
His	.66	≃ 1	.68	.68	.68	.77	≃ 1	.72	≃ 1	.68	≃ 1
Arg									
Trp									

Harbor seal His-Leu-Lys

All pinnipeds His-Leu-Lys

Sperm whale His-Leu-Lys

Peptide T-7, Residues #51-56. On occasion this peptide was found as two spots. This is consistent with the work of Dautrevaux et al., (1969) who found T-7 in two distinct peaks upon ion-exchange chromatography of peptide mixtures of bovine myoglobin. This is attributed to the formation of the sulfoxide of methionine. In the work reported here the analysis of the peptide spot with the higher R_f value usually displayed a methionine sulfoxide-sulfone peak equal in size to the methionine peak. The oxidation of methionine is the cause of the variable recovery of this amino acid for this peptide. The loss was excessive for the peptide from the myoglobins of Weddell seal and sperm whale. The highly variable value for alanine seen in the phocid analyses is attributed to free alanine. This could arise from position #90 from chymotryptic type cleavage or could be due to the peptide Ala-Ala-Lys (#143-145). The cleavage between Ile-Ala (#142-143) would produce such a peptide.

Harbor seal Ser-Glu-Asp-Asp-Met-Arg

Phocids Ser-Glx-Asx-Asx-Met-Arg

Walrus
Fur seal Ser-Glx-Asx-Glx-Met-Lys

Sperm whale Thr-Glu-Ala-Glu-Met-Lys

Peptide T-8, Residues #57-62. In the phocids, this peptide was usually poorly resolved from T-13, a peptide with 5 alanine residues. One map of Weddell seal myoglobin provided a fairly well resolved

Table 19 Amino acid residue composition of peptide T- 7 (Residues #51-56)

	Phocidae				Odobenidae	Otariidae	Cetacea	
AA	*Phoca* exp.	ref.	*Histrio-phoca* exp.	*Erig-nathus* exp.	*Leptony-chotes* exp.	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp.	ref.
Asp	2.37	2	2.38	1.93	1.62	1.29	1.00	.88	1
Thr								.83	
Ser	.74	1	.70	.93	.92	.69	.73	.65	1
Glu	1.23	1	1.28	1.41	1.38	2.30	2.24	2.80	3
Pro									
Gly									
Ala	1.04		2.00	.59	.46	.31		2.12	2
Val									
Met	.92	1	.79	.97	.42	.86	.76	.55	1
Ile									
Leu								.80	1
Tyr									
Phe									
Lys			.59	.35		1.10	1.10	2.07	1
His				.30					
Arg	.90	1	.82	.77	1.00				
Trp									

T-8 peptide, though the neighboring T-13 evidenced itself with an alanine contamination. The T-8 peptide from whale myoglobin contained a second lysine corresponding to position #63.

No peptides recognizable as T-8 were found in the fur seal and walrus myoglobin maps.

Harbor seal	Ser-Glu-Asp-Leu-Arg
Phocids	Ser-Glx-Asx-Leu-Arg
Sperm whale	Ala-Ser-Glu-Asp-Leu-Lys-Lys

Peptide T-9b, Residues #63-69. This peptide is produced by a chymotryptic type of cleavage. There is a difference between the sperm whale myoglobin sequence and that of the harbor seal. This difference persists in the phocids examined. In some phocid maps, there was incomplete resolution of this peptide from T-16 (Tyr-Lys). This was readily corrected for. The fur seal sequence is different from sperm whale and harbor seal by two replacements. A T-9b peptide was not found in the map of walrus myoglobin.

Harbor seal	His-Gly-Asp-Thr-Val-Leu
Phocids	His-Gly-Asx-Thr-Val-Leu
Fur seal	His-Gly-Gly-Ala-Val-Leu
Sperm whale	His-Gly-Val-Thr-Val-Leu

Peptide T-9c, Residues #70-77. This peptide results from a chymotryptic type of cleavage. The composition of this peptide in the phocids appears to be the same as that for the harbor seal. The corresponding peptide for fur seal and walrus myoglobin is different as evidenced by the fact that the analysis indicates a peptide two residues

Table 20 Amino acid residue composition of peptide T-8 (Residues #57-62)

	Phocidae					Odobenidae	Otariidae	Cetacea		
AA	*Phoca* exp*		ref.	*Histrio-phoca* exp.*	*Erig-nathus* exp.*	*Leptony-chotes* exp.	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp.	ref.
Asp	.96	≃	1	1.09	1.00	1.09			.97	≃ 1
Thr										
Ser	.59	≃	1	.81	.63	.67			.82	≃ 1
Glu	1.00	≃	1	1.03	1.00	1.39			1.19	≃ 1
Pro										
Gly						.36				
Ala						.66			1.03	≃ 1
Val										
Met										
Ile										
Leu	1.00	≃	1	1.08	1.00	.92			.88	≃ 1
Tyr										
Phe										
Lys										
His										
Arg	.83	≃	1	.99	.99	.82				
Trp										

*Data is derived from unresolved pair.

Table 21 Amino acid residue composition of peptide T-9b (Residues #64-67)

	Phocidae				Odobenidae	Otariidae	Cetacea
AA	*Phoca* exp. ≃ ref.	*Histriophoca* exp.	*Erignathus* exp.	*Leptonychotes* exp.	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp. ≃ ref.
Asp	.95 ≃ 1	.95	.89	.79		.39	
Thr	.72 ≃ 1	.71	.92	.90			1.00 ≃ 1
Ser							
Glu						1.11 ≃ 1	
Pro							
Gly	1.18 ≃ 1	1.22	1.16	1.21		1.01 ≃ 1	1.38 ≃ 1
Ala						1.06	
Val	.80 ≃ 1	.92	.89	.86		1.05 ≃ 1	2.28 ≃ 2
Met							
Ile							
Leu	1.27 ≃ 1	1.37	1.15	1.27		1.02 ≃ 1	1.33 ≃ 1
Tyr							
Phe							
Lys							
His	.97 ≃ 1	1.07	.56	.77		.82 ≃ 1	.64 ≃ 1
Arg							
Trp							

longer, (T-9e).

Harbor seal Thr-Ala-Leu-Gly-Gly-Ile-Leu-Lys

Phocids Thr-Ala-Leu-Gly-Gly-Ile-Leu-Lys

Sperm whale Thr-Ala-Leu-Gly-Ala-Ile-Leu-Lys

Peptide T-9e, Residues #68-77. This peptide was seen only in the myoglobins from fur seal and walrus. In the walrus, phenylalanine has been substituted for one of the leucines.

Harbor seal Val-Leu-Thr-Ala-Leu-Gly-Gly-Ile-Leu-Lys

Walrus Val(Leu)Thr-Ala(Leu)Gly-Gly-Ile(Phe)Lys

Fur seal Val-Leu-Thr-Ala-Leu-Gly-Gly-Ile-Leu-Lys

Sperm whale Val-Leu-Thr-Ala-Leu-Gly-Ala-Ile-Leu-Lys

Peptide T-10, series 10a and 10b, Residues #78-96. The analytical values obtained for this peptide are best attributed to a peptide containing three lysines, hence, 18 residues long. The sometimes marginal value for lysine is most probably due to the fact that one lysine residue is amino terminal in the peptide and subject to low recovery. It will be noted that there is a "family" of T-10 peptides in the same area of the map. Recognition of the center of one spot relative to the center of a neighboring spot is difficult since large peptides do not produce equivalent amounts of color, and relative amounts of neighboring peptides may vary. The exact composition of the material eluted from the peptide-containing paper may reflect the wisdom of a given decision when cutting the peptide spot paper. Particularly when peptides are close to one another. Neither T-10a nor T-10b were found in the sperm whale maps.

Table 22 Amino acid residue composition of peptide T- 9c (Residues #68-77)
9e (Residues #66-77)

	Phocidae				Odobenidae	Otariidae	Cetacea	
AA	*Phoca* exp.	ref.	*Histrio-phoca* exp.	*Erig-nathus* exp.	*Leptony-chotes* exp.	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp. ref.
Asp								
Thr	.64	≃ 1	.70	.79	.64	1.33 ≃ 1	1.87 ≃ 2	.96 ≃ 1
Ser								
Glu	.32		.31	.33	.34	.54		
Pro								
Gly	2.10	≃ 2	2.08	2.12	1.96	1.90 ≃ 2	1.95 ≃ 2	1.83 ≃ 2
Ala	1.00	≃ 1	.98	1.05	1.05	1.20 ≃ 1	1.08 ≃ 1	2.25 ≃ 2
Val								
Met								
Ile	.92	≃ 1	.90	.97	.95	.78 ≃ 1	1.03 ≃ 1	.92 ≃ 1
Leu	2.33	≃ 2	2.36	2.38	2.44	2.34 ≃ 2	3.12 ≃ 3	1.79 ≃ 2
Tyr								
Phe	.35			.31	.38	1.07 ≃ 1		
Lys	.82	≃ 1	.85	.96	.82	.77 ≃ 1	.96 ≃ 1	1.00 ≃ 1
His								
Arg								
Trp								

Table 23 Amino acid residue composition of peptide T- 10a (Residues #79-96)
10b (Residues #78-96)

	Phocidae					Odobenidae	Otariidae	Cetacea	
AA	*Phoca* exp.	ref.	*Histrio-phoca* exp.	*Erig-nathus* exp.	*Leptony-chotes* exp.	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp.	ref.
Asp					.32				
Thr	1.14	≃ 1	.86		1.09				
Ser	1.30	≃ 1	.91		1.09				
Glu	3.18	≃ 3	3.09		3.07				
Pro	.89	≃ 1	1.04		.57				
Gly	1.00	≃ 1	.91		.91				
Ala	3.34	≃ 3	3.14		2.98				
Val									
Met									
Ile									
Leu	1.91	≃ 2	1.86		2.00				
Tyr									
Phe									
Lys	2.62	≃ 3	2.22		2.61				
His	2.91	≃ 3	2.94		2.89				
Arg									
Trp									

A clear-cut T-10a or T-10b was not found on the bearded seal maps. There were spots in the proper areas for these peptides, but the recovery was too low to permit definitive analyses. Qualitatively, the analyses resembled the T-10 series. T-10d was found for bearded seal, suggesting that the problem encountered for T-10a or T-10b was not due to a marked difference in the peptide sequence.

Harbor seal	Lys-Lys-Gly-His-His-Glu-Ala-Glu-Leu-Lys-Pro-Leu Ala-Gln-Ser-His-Ala-Thr-Lys
Harbor, ribbon, and bearded seal	Lys-Lys-Gly-His-His-Glx-Ala-Glx-Leu-Lys-Pro-Leu Ala-Glx-Ser-His-Ala-Thr-Lys
Sperm whale	Lys-Lys-Gly-His-His-Glu-Ala-Glu-Leu-Lys-Pro-Leu Ala-Gln-Gln-Ser-His-Ala-Thr-Lys

Peptide T-10d, Residues #90-96. This "sub-peptide" of T-10a is helpful in the alignment by homology for this long peptide sequence and is considered identical for all species examined. This peptide was not found in maps of Weddell seal myoglobin. Since T-10a was determined for this species, the T-10d peptide was not pursued when duplication failed to recover it.

Harbor seal	Ala-Glu-Ser-His-Ala-Thr-Lys
All pinnipeds	Ala-Glx-Ser-His-Ala-Thr-Lys
Sperm whale	Ala-Gln-Ser-His-Ala-Thr-Lys

Peptide T-11, Residues #97-102. This peptide showed the same composition for all species examined. It is concluded from this data, that the sequence for this peptide is identical in all species examined.

Table 24 Amino acid residue composition of peptide T- 10d (Residues #90-96)

	Phocidae				Odobenidae	Otariidae	Cetacea	
AA	*Phoca* exp.	ref.	*Histrio-phoca* exp.	*Erig-nathus* exp.	*Leptony-chotes* exp.	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp. ref.
Asp						.38	.39	
Thr	.70	≃ 1	.80	.88		.70 ≃ 1	.74 ≃ 1	.72 ≃ 1
Ser	.77	≃ 1	.84	.87		.70 ≃ 1	.75 ≃ 1	.88 ≃ 1
Glu	1.33	≃ 1	1.10	1.11		1.13 ≃ 1	1.10 ≃ 1	1.36 ≃ 1
Pro								
Gly								
Ala	1.68	≃ 2	1.80	1.73		1.69 ≃ 2	1.80 ≃ 2	1.72 ≃ 2
Val								
Met								
Ile								
Leu	.54		.35			.38	.31	
Tyr								
Phe								
Lys	1.00	≃ 1	.98	1.11		1.40 ≃ 1	1.18 ≃ 1	1.36 ≃ 1
His	1.04	≃ 1	1.04	1.07		1.08 ≃ 1	1.06 ≃ 1	.96 ≃ 1
Arg								
Trp								

Table 25 Amino acid residue composition of peptide T-11 (Residues #97-102)

	Phocidae					Odobenidae	Otariidae	Cetacea	
AA	*Phoca* exp.	ref.	*Histrio-phoca* exp.	*Erig-nathus* exp.	*Leptony-chotes* exp.	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp.	ref.
Asp									
Thr									
Ser									
Glu									
Pro	.85	≃ 1	1.02	.98	.71	1.05 ≃ 1	.92 ≃ 1	1.02	≃ 1
Gly									
Ala									
Val									
Met									
Ile	2.35	≃ 2	1.98	2.15	2.32	1.78 ≃ 2	1.91 ≃ 2	2.10	≃ 2
Leu									
Tyr									
Phe									
Lys	2.06	≃ 2	1.83	1.96	2.18	2.14 ≃ 2	2.01 ≃ 2	2.19	≃ 2
His	.75	≃ 1	.66	.75	.77	.85 ≃ 1	.69 ≃ 1	.69	≃ 1
Arg									
Trp									

Harbor seal His-Lys-Ile-Pro-Ile-Lys

All pinnipeds His-Lys-Ile-Pro-Ile-Lys

Sperm whale His-Lys-Ile-Pro-Ile-Lys

Peptide T-11b, Residues #99-102. This short peptide was not recovered from all phocid maps. It was recovered from walrus and fur seal and helped establish identity in this area of the sequence.

Harbor seal Ile-Pro-Ile-Lys

Walrus & fur seal Ile-Pro-Ile-Lys

Sperm whale Ile-Pro-Ile-Lys

Peptide T-12a, Residues #103-106. This peptide provided poor data. The low recovery for the tyrosine reported elsewhere was extreme here and may be due to contamination with the "sub-peptide" (Leu-Glu-Phe).

In the peptide maps of walrus myoglobin, there was incomplete resolution of this peptide from T-1d. As mentioned in the discussion of T-1d, it is possible to "extract" T-12a from the analytical results, however.

Harbor seal Tyr-Leu-Glu-Phe

All pinnipeds Tyr-Leu-Glx-Phe

Sperm whale Tyr-Leu-Glu-Phe

Peptide T-12b, Residues #107-115. This peptide was not found in any of the maps.

Peptide T-12c, Residues #116-118. This peptide was variably contaminated by the neighboring free amino acid lysine. The peptide from the myoglobin of fur seal and walrus occupied the same area as the

Table 26 Amino acid residue composition of peptide T-11b (Residues #99-102)

	Phocidae				Odobenidae	Otariidae	Cetacea	
AA	*Phoca* exp.	ref.	*Histrio-phoca* exp.	*Erig-nathus* exp.	*Leptony-chotes* exp.	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp. ref.
Asp								
Thr								
Ser					.73	1.11 ≃ 1	.95 ≃ 1	
Glu								
Pro								
Gly								
Ala								
Val								
Met								
Ile					2.22	1.73 ≃ 1	1.74 ≃ 2	
Leu								
Tyr								
Phe								
Lys					1.24	1.11 ≃ 1	1.04 ≃ 1	
His							.38	
Arg							.34	
Trp								

Table 27 Amino acid residue composition of peptide T-12a (Residues #103-106)

	Phocidae					Odobenidae		Otariidae	Cetacea	
AA	*Phoca* exp.	ref.	*Histriophoca* exp.	*Erignathus* exp.	*Leptonychotes* exp.	*Odobenus* exp.		*Callorhinus* exp.	*Physeter* exp.	ref.
Asp										
Thr										
Ser									.35	
Glu	.82	≃ 1	.81	.82	.80	.88	≃ 1		1.00	≃ 1
Pro										
Gly	.42		.48		.39				.56	
Ala	.47				.37				.71	
Val										
Met										
Ile										
Leu	1.60	≃ 1	1.40	1.17	1.27	.86	≃ 1		1.23	≃ 1
Tyr	t	≃ 1	.13	.36	.39	.42	≃ 1		.15	≃ 1
Phe	.61	≃ 1	.61	.77	.69	1.08	≃ 1		.35	≃ 1
Lys										
His										
Arg										
Trp										

Table 28 Amino acid residue composition of peptide T- 12c (Residues #116-118)

	Phocidae				Odobenidae	Otariidae	Cetacea	
AA	*Phoca* exp.	ref.	*Histrio-phoca* exp.	*Erig-nathus* exp.	*Leptony-chotes* exp.	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp. ref.
Asp								
Thr								
Ser	.75	≃ 1	.75	.90	.75			1.17 ≃ 1
Glu								
Pro								
Gly			.45			1.05 ≃ 1	.99 ≃ 1	
Ala								
Val								
Met								
Ile								
Leu								
Tyr								
Phe								
Lys	1.32	≃ 1	2.45	1.19	2.10	2.08 ≃ 1	1.26 ≃ 1	1.73
His	.60	≃ 1	1.01	.61	1.05	.67 ≃ 1	.64 ≃ 1	1.00 ≃ 1
Arg								.83 ≃ 1
Trp								

corresponding peptide in the phocids, but analyses indicated a glycine for serine substitution. This finding was reproducible.

Harbor seal	His-Ser-Lys
Phocids	His-Ser-Lys
Walrus & fur seal	His-Gly-Lys
Sperm whale	His-Ser-Lys

Peptide T-13, Residues #119-132. In the phocids, this peptide was usually poorly resolved from T-8. As a result, these two peptides are only provisionally identified. This peptide was not found in the maps of bearded seal, walrus, Weddell seal, and fur seal myoglobin.

Harbor seal	His-Pro-Ala-Gln-Phe-Gly-Ala-Asp-Ala-Gln-Ala-Ala-Met-Lys
Phocids (2)	His-Pro-Ala-Gln-Phe-Gly-Ala-Asp-Ala-Gln-Ala-Ala-Met-Lys
Sperm whale	His-Pro-Gly-Asn-Phe-Gly-Ala-Asp-Ala-Gln-Gly-Ala-Met-Asp-Lys

Peptide T-14, Residues #134-139. The data for this peptide indicates that this portion of the molecule is identical for all species examined.

Harbor seal	Ala-Leu-Glu-Leu-Phe-Arg
All pinnipeds	Ala-Leu-Glx-Leu-Phe-Arg
Sperm whale	Ala-Leu-Glu-Leu-Phe-Arg

Peptide T-15, Residues #140-145. This peptide is unique in composition and readily identified in its mixture with T-4 in the phocids.

Table 29 Amino acid residue composition of peptide T- 13 (Residues #119-132)

	Phocidae					Odobenidae	Otariidae	Cetacea	
AA	*Phoca* exp.*	ref.	*Histrio-phoca* exp.*	*Erig-nathus* exp.	*Leptony-chotes* exp.	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp.	ref.
Asp	.96	≃ 1	1.09					2.74	≃ 3
Thr									
Ser									
Glu	1.93	≃ 2	2.08					1.33	≃ 1
Pro	.70	≃ 1	1.11					.78	≃ 1
Gly	1.22	≃ 1	1.15					2.85	≃ 3
Ala	4.80	≃ 5	4.72					2.90	≃ 3
Val									
Met	.66	≃ 1	.68					.46	≃ 1
Ile									
Leu									
Tyr									
Phe	.95	≃ 1	.88					1.00	≃ 1
Lys	1.03	≃ 1	.99					1.27	≃ 1
His	.69	≃ 1	.65					.85	≃ 1
Arg									
Trp									

* Data is derived from unresolved pair.

Table 30 Amino acid residue composition of peptide T-14 (Residues #134-139)

	Phocidae				Odobenidae	Otariidae	Cetacea	
AA	*Phoca* exp.	ref.	*Histrio-phoca* exp.	*Erig-nathus* exp.	*Leptony-chotes* exp.	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp.	ref.
Asp									
Thr	.38								
Ser									
Glu	.86	≃ 1	.91	.83	.95	.95	≃ 1	1.15	≃ 1	.81	≃ 1
Pro									
Gly	.55		.56	.56	.44	.42		.32		.78	
Ala	.68	≃ 1	.85	.73	.64	.92	≃ 1	.91	≃ 1	1.56	≃ 1
Val									
Met									
Ile									
Leu	2.58	≃ 2	2.36	2.10	2.20	2.05	≃ 2	1.93	≃ 2	2.44	≃ 2
Tyr									
Phe	.82	≃ 1	.95	.58	.97	1.07	≃ 1	1.01	≃ 1	.81	≃ 1
Lys									
His									
Arg	.87	≃ 1	.86	.74	.84	1.01	≃ 1	1.00	≃ 1	.80	≃ 1
Trp									

Table 31 Amino acid residue composition of peptide T- 15 (Residues #140-145)

	Phocidae				Odobenidae	Otariidae	Cetacea	
AA	*Phoca* exp.*	ref.	*Histrio-phoca* exp.	*Erig-nathus* exp.*	*Leptony-chotes* exp.*	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp.	ref.
Asp Thr Ser	2.18 ≃	2	1.85	1.65	1.85	1.74 ≃ 2	1.81 ≃ 2	.59 ≃	1
Glu Pro Gly									
Ala Val Met	2.30 ≃	2	2.35	2.70	2.06	2.16 ≃ 2	2.20 ≃ 2	1.94 ≃	2
Ile Leu Tyr	.79 ≃	1	.64	.76	.79	.79 ≃ 1	.75 ≃ 1	.91 ≃	1
Phe Lys His	1.00 ≃	1	.97	1.24	.99	1.21 ≃ 1	1.19 ≃ 1	1.05 ≃	1
Arg Trp									

*Data derived from unresolved pairs.

Harbor seal Asn-Asp-Ile-Ala-Ala-Lys

All pinnipeds Asx-Asx-Ile-Ala-Ala-Lys

Sperm whale Asp-Ile-Ala-Ala-Lys

Peptide T-16, Residues #146-147. The low recovery of tyrosine has been treated elsewhere. The Tyr-Arg peptide found in walrus and fur seal is poorly resolved from T-5b (Phe-Lys). The best values were obtained from the walrus map.

Harbor seal Tyr-Lys

Phocids Tyr-Lys

Walrus
fur seal Tyr-Arg

Sperm whale Tyr-Lys

Peptide T-17, Residues #148-153. This is the carboxyl terminal peptide of the myoglobin. The composition of this peptide in phocid myoglobins was identical. The walrus and fur seal myoglobins are seen to stand somewhere between the phocids and the sperm whale in this peptide. Their myoglobins possess the phenylalanine of the phocid myoglobins but the glutamine of the sperm whale myoglobin.

Harbor seal Glu-Leu-Gly-Phe-His-Gly

Phocids Glx-Leu-Gly-Phe-His-Gly

Walrus, fur seal Glx-Leu-Gly-Phe-Glx-Gly

Sperm whale Glu-Leu-Gly-Tyr-Gln-Gly

Myoglobin Sequence

It is now possible to assemble the individual peptides into the larger parent myoglobin molecule. Comparative sequences are given for the myoglobin of the phocid seals, the northern fur seal and the walrus

Table 32 Amino acid residue composition of peptide T- 16 (Residues #146-147)

	Phocidae					Odobenidae		Otariidae		Cetacea	
	Phoca		*Histrio-phoca*	*Erig-nathus*	*Leptony-chotes*	*Odobenus*		*Callorhinus*		*Physeter*	
AA	exp.	ref.	exp.	exp.	exp.	exp.		exp.		exp.	ref.
Asp											
Thr											
Ser											
Glu											
Pro											
Gly											
Ala											
Val											
Met											
Ile											
Leu	.32		.32	.34							
Tyr	.42	≃ 1	.42	.45	.46	.62	1	.66	≃ 1	.42	≃ 1
Phe								.58			
Lys	1.00	≃ 1	1.00	1.00	1.00	.34		.72		1.00	≃ 1
His											
Arg						.93	1	.88	≃ 1		
Trp											

Table 33 Amino acid residue composition of peptide T- 17 (Residues #148-153)

	Phocidae				Odobenidae	Otariidae	Cetacea	
AA	*Phoca* exp.	ref.	*Histrio-phoca* exp.	*Erig-nathus* exp.	*Leptony-chotes* exp.	*Odobenus* exp.	*Callorhinus* exp.	*Physeter* exp.	ref.
Asp									
Thr									
Ser									
Glu	.72	≃ 1	.80	.76	.81	1.86 ≃ 2	1.78 ≃ 2	2.29	≃ 2
Pro									
Gly	2.23	≃ 2	2.00	2.24	2.30	2.08 ≃ 2	1.97 ≃ 2	2.36	≃ 2
Ala									
Val									
Met									
Ile									
Leu	1.19	≃ 1	1.17	1.14	1.18	1.07 ≃ 1	1.15 ≃ 1	.82	≃ 1
Tyr								.56	≃ 1
Phe	.98	≃ 1	.86	.97	1.00	.78 ≃ 1	.89 ≃ 1		
Lys									
His	1.06	≃ 1	.91	1.01	.98				
Arg									
Trp									

in figures 13, 14, and 15. These sequences are given in terms of the one letter code of the amino acids (Dayhoff 1969). The sequence for harbor seal myoglobin (Bradshaw and Gurd 1969) is employed as a base. The uncertainty inherent in the placement of glutamic acid-glutamine and aspartic acid-asparagine is indicated by the use of "Z" and "B" respectively. An entry is made only if there is a difference in the sequence given. The dashed lines within brackets indicate peptides for which no information was obtained.

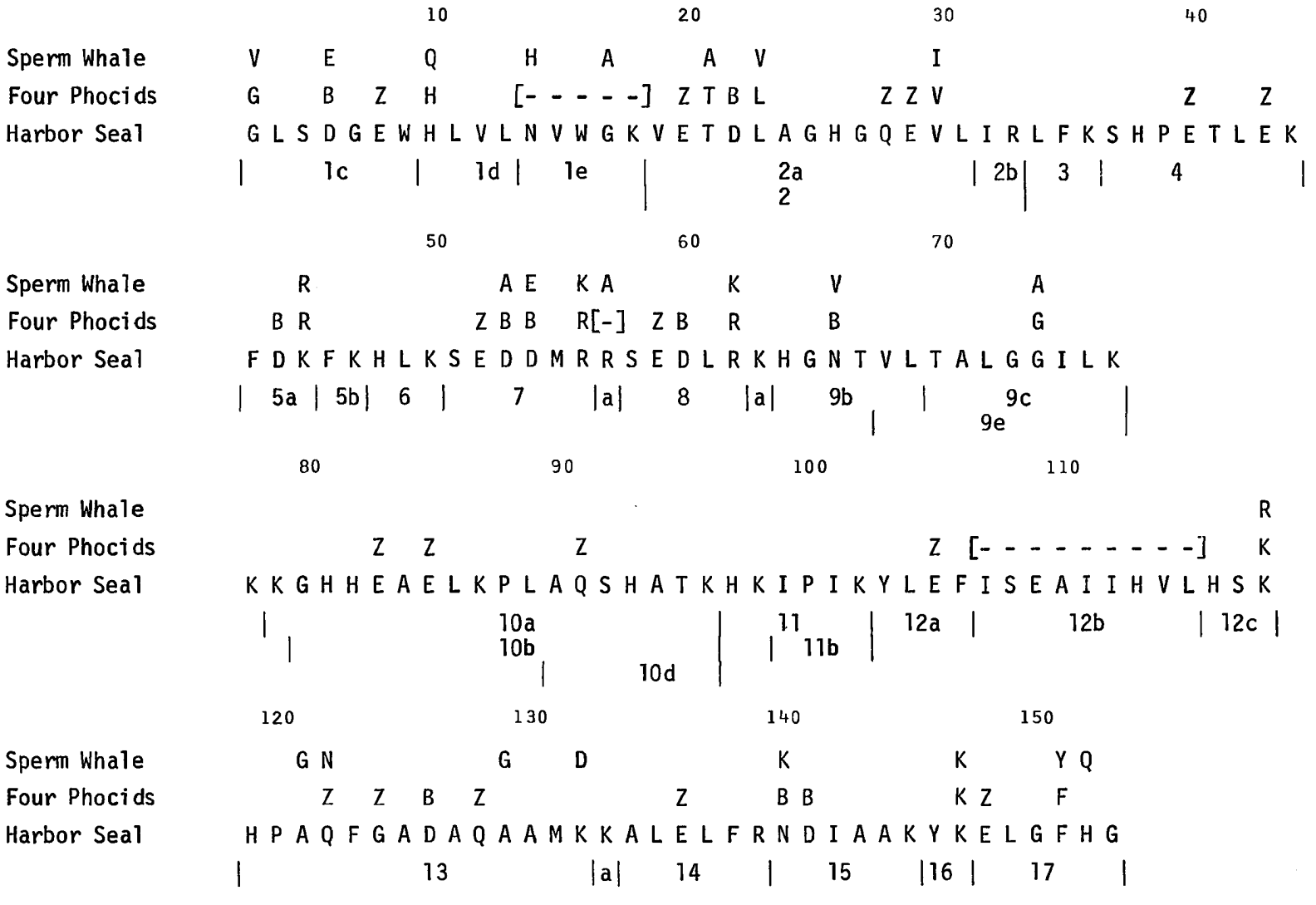

Figure 12. Suggested amino acid sequence of the myoglobin of the four phocids: Bering Sea harbor seal, ribbon seal, bearded seal, and Weddell seal. T-13 was not found for bearded seal and Weddell seal. Parenthesized dashes indicate sequences for which no information was obtained.

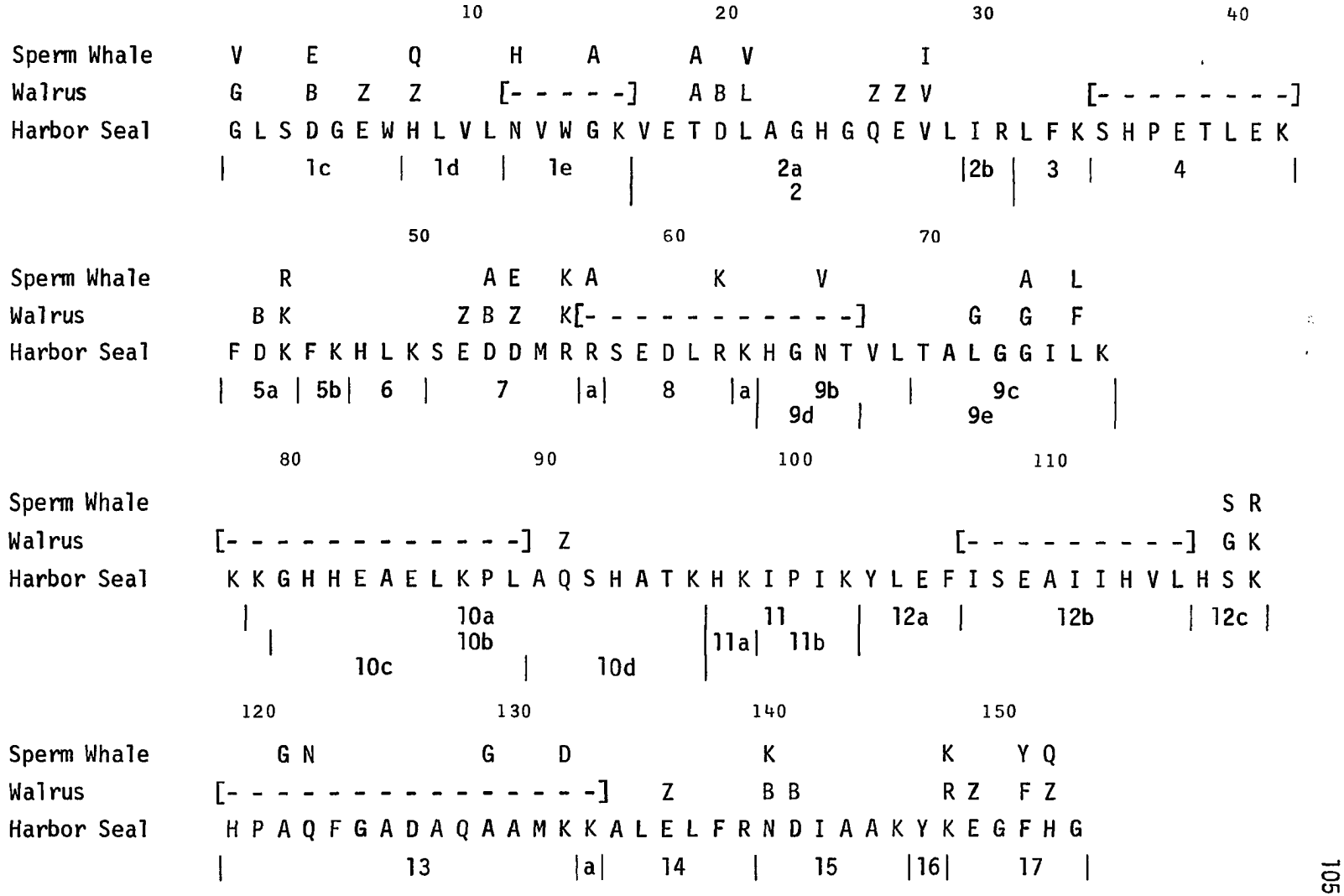

Figure 13. Suggested amino acid sequence of the myoglobin of the walrus. Parenthesized dashes indicate sequences for which no information was obtained.

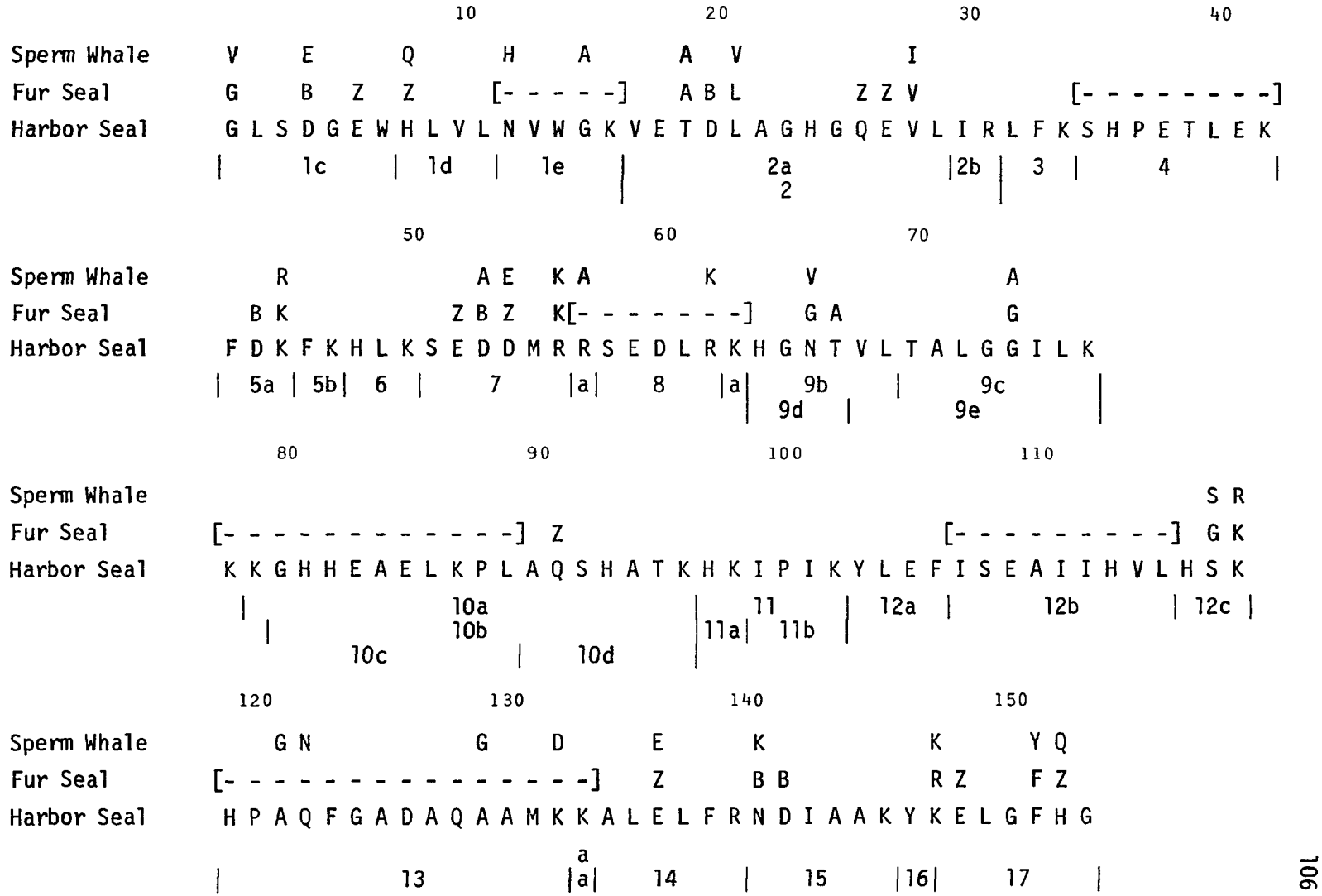

Figure 14. Suggested amino sequence of the myoglobin of the northern fur seal. Parenthesized dashes indicate sequences for which no data was obtained.

DISCUSSION

The fossil record of the pinnipeds begins with forms already highly adapted and specialized (Romer, 1962). They are generally considered to have evolved from creodont ancestors which were prevalent during the Paleocene, 75 million years ago. A branch of the creodonts, the Miacidae, survived their creodont progenitors and at the end of the Eocene, some 60 million years ago, gave rise to the three major branches of the Carnivora: the Feloidea, the Canoidea, and the Pinnipedia. The precise rank and relationship of the pinnipeds to other groups is not settled. Recent work suggests that the pinnipeds should be placed within the Canoidea as a suborder (Sarich, 1969a). This conclusion was drawn from the results of immunological studies based on albumin as the protein marker. No light can be shed on this problem by the study reported here.

The relationships and ranks within the pinniped order or suborder are also considered open to question and are presently being studied in some detail (Sarich, 1969a, 1969b; Burns and Fay, 1970). There are three recognized families within the pinnipedia. Family Otariidae comprises the fur seals and sea lions. Family Odobenidae is composed of the Pacific and the Atlantic walruses. Family Phocidae comprises the "earless" seals. The phocid family is further divided into three tribes (Scheffer, 1958). The Phocini tribe is represented in this work by the harbor, and ribbon seals; tribe Erignathini is represented by the bearded seal; the Lobodoni tribe is represented by the Weddell seal, the single Antarctic species studied here.

When considering the systematic or taxonomic significance of the results of this study, it is necessary to place amino acid replacements in proper perspective. There are two important criteria which should be met if structural differences in homologous proteins are to be used as relevent homologous characters. First, protein differences must be inherited and relatively free of environmental effects; and second, adequate function of the chosen protein must be essential to the survival of the individual or population. The satisfaction of each criterion will be treated in turn.

Current concepts of protein synthesis are based on a wealth of material confirming the first criterion. The chosen model assumes that genetic information resides in the base sequences in the DNA molecule which in turn directly determine the amino acid sequence in each protein. Herein lies the great value of protein sequence for comparative work in the area of systematics. A taxonomic hierarchy is often based on characters that do not necessarily reflect the genetic constitution of the organism in question. Parasite load, availability, and quality of food at critical times during development have significant effect on such parameters as general size, bone length, skull dimensions, etc. Amino acid sequence studies, however, provide unit differences that are directly related to the genetic constitution of the organism in question by counting the substitutions of one amino acid for another.

As is well established for microorganisms, most amino acids are coded for by one or more base triplets (Cold Spring Harbor Symposium, 1966). Although methionine and tryptophan each are coded by a single

triplet, and redundancy in the genetic code can be as high as six for serine, arginine and leucine. Such high levels of redundancy may be considered protective, however, since they permit a certain amount of mutation or change in the DNA base sequence without causing changes in the corresponding protein.

There is, then, the distinct possibility that a base change in DNA can occur without an observed effect on the primary structure of the protein whose controlling base sequence has been altered. However, on a statistical basis such silent mutations probably occur less frequently than those yielding obvious changes in protein sequence. Therefore, identity of structure in homologous proteins in two individuals is assumed to show greater genetic similarity than non-identity.

The method of separation and analysis used in this study provides data for most of the myoglobin molecule from the phocids studied here. While the same degree of success was not achieved in the myoglobins of the representatives of the other two families, the information obtained for the latter does provide some basis for comparisons. Because the analytical method used could not distinguish glutamic acid from glutamine, the presence of the former in a peptide was interpreted as indicating whichever of the two was found in the homologous position of the comparison standard. Aspartic acid and asparagine were treated similarly. This assumption of identity is statistically warranted since maintenance of identity is more probable than change in such situations. In the following, the differences observed are treated as genetic differences. The question of function raised above and the chemical significance of

the replacements will be treated in a separate section.

Comparison of the Sequence of Myoglobins from Geographically Distinct Subspecies of the Harbor Seal

One observation possible from the work described here is that the isolated tryptic peptides from the myoglobin of Bering Sea harbor seal may be aligned directly with the sequence of myoglobin from harbor seal from the coast of Maine (Bradshaw et al. 1969b). This is interesting since one would expect populations so far removed from one another geographically to display some genetic diversity. It is quite possible that the two populations of harbor seal under consideration have not been isolated from one another for a long enough period of time for mutations to have become established. Since the myoglobins of all four phocids examined appear to be identical, the identity of myoglobins isolated from distinct populations of one phocid species is of less significance.

Comparison of Sequences of the Myoglobins Isolated from Members of Family Phocidae

The sequence of the myoglobins of harbor seal, ribbon seal, bearded seal, and Weddell seal appear to be identical. Peptides T-1e and T-12b were not found, but for the purpose of discussion, the lack of these two peptides will be ignored, as will the fact that peptide T-13 of Weddell seal and bearded seal myoglobins did not provide satisfactory analyses.

The apparent identity of myoglobins from these four phocids is consistent with the work of Fay et al. (1967) who found greater karyotypic similarity among four (harbor, ringed, bearded, and Weddell seals)

members of the phocidae than to that of species of the other two pinniped families. It is also consistent with Sarich's (1969b) recent work employing the immunology of albumin as a marker.

The fact that no difference is seen in the sequences of myoglobins of these four representatives of the family phocidae suggests that myoglobin may be of some use in defining family (or subfamily) rank, because the myoglobins of the other two families are distinctly different.

Comparison of the Sequences of Myoglobins from the Three Families

The myoglobins of walrus and fur seal presented a little more difficulty in recognition of peptides. In fact there were a few peptide spots whose analyses could not be satisfactorily aligned with the known harbor seal sequence. Even so, these two proteins did provide peptide maps resembling one another very closely. The peptide analyses for these two proteins resemble duplicates from the same protein. Final comparisons suggest that these two myoglobins are almost identical in the portions of the molecule identified. The exception is a phenylalanine for leucine substitution in Peptide T-9b.

Eight replacement differences were found between the sequences of phocid myoglobins and that of the fur seal. Nine differences were seen in the case of the walrus. The failure to obtain clear-cut analyses for peptides T-4 and T-13 in the fur seal and walrus suggests that there are even more differences than those found.

Quite obviously, if one were to use myoglobin sequence as a marker, then the pinnipeds should be divided into two subgroups: the phocids

in one, and the walrus in the group containing the otariids. This assumes that the other otariids will display the same constancy in their myoglobin sequence as the phocids examined here. This similarity of the myoglobins of fur seal and walrus and their difference from myoglobins of representative phocids is consistent with the immunological work reported by Sarich (1969b). He found that the pinniped albumins could be classified as two groups: one, the phocids, and the other the otariids plus the walrus.

This rather distinct dichotomy of the pinnipeds may also be used to support the contention that these two groups arose separately and resemble one another due to convergence (McLaren, 1960). Comparison of the sequence of myoglobins from the pinnipeds with sequences of the myoglobins from representative members of the previously mentioned Canidae, Ursidae, Procyonidae, and Mustelidae would assist in the resolution of this problem. Such data is not presently available. Sarich's immunological distances indicate that the pinnipeds are of monophyletic origin, but he did note that confirmation would have to come from a second protein. The use of protein sequence as confirmation of immunological findings may be superior to a second series of immunological testings with a second antigen.

Comparison of myoglobins of pinnipeds with that of the sperm whale. Table 34 lists the differences seen between the myoglobin of the sperm whale and that of the pinnipeds studied here. The number of differences observed is of the same order of magnitude. This is consistent with a monophyletic origin of the pinnipeds. Nine of the differences seen are

Table 34. Comparison of varient residues in the Myoglobins of the pinnipeds and the sperm whale

Position No.	Harbor seal	Walrus	Fur seal	Sperm whale	Sperm whale vs. Phocids	Sperm whale vs. Walrus	Sperm whale vs. Fur seal
1	Gly	Gly	Gly	Val	1	1	1
4	Asp	Asx	Asx	Glu	1	1	1
8	His	Glx	Glx	Glu	1	0	0
19	Thr	Ala	-	Ala	1	0	0
21	Leu	Leu	Leu	Val	1	1	1
27	Glu	Glx	Glx	Asp	1	1	1
35	Ser	Gly	-	Ser	0	1	-
39	Thr	Ala	-	Thr	0	1	-
45	Lys	Lys	Lys	Arg	1	1	1
53	Asp	Asx	Asx	Ala	1	1	1
54	Asp	Asx	Glx	Glu	1	1	0
56	Arg	Lys	Lys	Lys	1	0	0
62	Arg	-	-	Lys	1	-	-
65	Asp	-	Gly	Val	1	-	1
74	Gly	Gly	Gly	Ala	1	1	1
76	Leu	Phe	Leu	Leu	0	1	0
117	Ser	Gly	Gly	Ser	0	1	1
121	Ala	-	-	Gly	1	-	-
140	Asn	Asx	Asx	Lys	1	1	1
147	Lys	Arg	Arg	Lys	0	1	1
151	Phe	Phe	Phe	Tyr	1	1	1
152	His	Glx	Glx	Gln	1	0	0
					17	15	12

common to the three families providing information for an ancestral pinniped myoglobin.

The Biochemical Significance of Variation of the Primary Structure of Hemoglobins and Myoglobins

The fact of protein variation both within and between taxonomic groups is now well established; what has yet to be determined are the basic mechanisms or principles involved which allow such variation. While the initial genetic change presumably results from random mutation in the chromosomal DNA, the reason for the persistence of any given change in the protein molecule must reside, somehow, in that molecule, or some circumstance surrounding its function or action.

It is customary to think of the primary structure of a protein as the product of a particular polynucleotide sequence in DNA. However, it is evident that the actual persistence of a given amino acid sequence in turn fixes the DNA sequence for the population or the species involved. If the protein sequence produced fails to perform the function or action required, and that function or action is important to survival, it is expected that the organism involved will be at a selective disadvantage and this particular amino acid sequence will not be perpetuated as typical for the species.

Protein studies may also reveal principles governing the types of base changes thought to be responsible for the amino acid replacements seen. An analysis of the types of base changes required to effect the amino acid replacements observed in human hemoglobin and in cytochrome c from other animals and plants provides two different patterns. In hemoglobin, most replacements are explainable as the

result of single purine to purine base changes (Vogel, 1969). In cytochrome c, the replacements reflect pyrimidine-purine changes in the first two codon positions. The difference perhaps may be attributed to differences in the tertiary structure of these compounds. The two globins possess an internally ordered structure based on the alpha helix. However, cytochrome c does not exhibit a significant percentage of helical zone, indicating a lack of regular internal order. Since the mechanism that produces base changes in the DNA molecule is assumed to be the same for the total length of all chromosomes, then the mechanism that genetically fixes a given amino acid in the protein molecule must not operate at the DNA level. This suggests that the mechanism involved operates at the level of the protein itself. It also suggests that each homologous protein series should be considered separately, until such a time that a clear view is gained.

On the assumption that an amino acid replacement by a very similar amino acid would be more likely to persist than one by a dissimilar, since the former would alter the protein the least, Sneath (1965) considered a number of characteristics of the amino acids. He included some thirty-two different parameters, from molecular weight of the side chain to the color produced by reaction on paper with ninhydrin when collidine is included in the reagent. A characteristic number for each amino acid was derived from weighted values of these variables using a formula developed earlier by Sneath and Sokol (1963). Differences between these values for any given pair is an index of their similarity, with the value of one indicating complete identity. The

treatment so provided may well be the first step in the right direction, but it failed to take proper note of the fact that most proteins exist in a three-dimensional aqueous environment.

For hemoglobin and myoglobin, a few patterns are seen that indicate that the environment at the sequence position is a major factor in the persistence of a given kind of amino acid residue. Perutz et al. (1965) points out that there is a common factor found in hemoglobin and myoglobin. The non-polar residues are located mainly in the interior of the globin chain. To quote directly "Long alpha helical segments exhibit a regular periodicity of invariant non-polar sites". (op. cit. p. 670). The work of Perutz et al. (1965) and Edmundson (1965) suggests that the major factors in stabilizing the structure of these globins are hydrophilic and hydrophobic bonding. Dickerson and Geis (1969) have recently published a general treatment of the information presently available on the subject of the three-dimensional structure of proteins, in which Dickerson refers to the "hydrophobic-in, hydrophilic-out" principle.

The development of this principle parallels the increase in the quantity of data concerning the tertiary structure of the globins. Kendrew et al. (1961) noted that in crysta-line sperm whale myoglobin, 30% of the polar residues were directed internally, and 45% of the non-polar were directed externally. In particular, all phenylalanine and methionine residues were directed internally, and almost all lysine, arginine, and glutamic acid residues were externally directed. On the other hand, almost all alanines were directed externally.

A particularly interesting statment made by Perutz (1965), emphasizes the possible general applicability of the concept of differences in orientation of polar and non-polar residues.

> The interior of the sperm whale myoglobin molecule is made up of non-polar residues, almost everywhere in Van der Waals contact with their neighbors. Tryptophan and tyrosine are considered non-polar, the large, rigid carbon skeleton of which causes them to seek a non-polar environment. In hemoglobin, the glycines and alanines being only weakly hydrophobic, lie on the surface. Large non-polar side chains which are not interior tend to bury themselves in crevices near the surface, offering minimum contact with the water.

The following model is an attempt to bridge the gap between studies of acceptability of amino acid replacements and the concepts of Kendrew and Perutz.

As a first approximation, consider that the polypeptide chain is composed of a series of peptide bonds creating a glycine backbone. To this glycine polymer, there are attached, at the alpha carbons, certain substituents, the amino acid side chains.

Since the globins are in an aqueous environment during their functional lifetime, the change in the tendency of each externally directed side chain groups to form hydrogen bonds with the surrounding water molecules may be taken as an indicator of the chemical change imposed by a particular amino acid replacement at an external site. Similarly, the change in tendency of internally directed side chain groups to form hydrophobic bonds can also be taken as an indicator of the significance of replacements of internally directed groups. Hence, the problem is reduced to a search for an index of the polarity of the substituent side groups involved. The dipole moment appears

to be satisfactory in this respect, and the system proposed here is based on the published dipole moment values for the side chain groups (McClellan, 1963).

There is considerable agreement between an intuitive grouping of polarity types of the amino acids and the dipole moments of the amino acid sidechains (Table 35). Cysteine as a source of covalent inter- and intra-molecular bonding and proline as a helix interrupter must receive special treatment.

With these reservations in mind, a matrix based on the absolute differences between the dipole moments of the amino acids can be constructured (Table 36). This dissimilarity matrix is an index of the degree of chemical change imposed on the position in question by a given substitution. A glycine for alanine substitution is trivial on this basis, whereas, a glycine for glutamine substitution (Δ=3.4) is accorded some importance. This matrix could provide the basis for a numerical treatment of amino acid replacements.

The principle of isopolar substitution is attractive when first considering amino acid replacements. However, that it is not the major factor in the persistence of substituents in human hemoglobin is shown in Table 37 where 41 of the 59 replacements listed are not isopolar. There are at least two other factors. First, polar-non-polar substitutions might be beneficial in positions where such changes stabilize the three dimensional structure of the molecule; and second, relative polarity might be a negligible factor compared to chemical properities in locations intimately involved in the function of the molecule.

Table 35

Dipole Moments of Amino Acid Side Chains

Small aliphatic	
Glycine	0
Alanine	0
Aliphatic	
Methionine	0
Valine	0
Leucine	0
Isoleucine	.13
Hydroxyl group	
Serine	1.7
Threonine	1.7
Tyrosine	1.6
Aromatic	
Phenylalanine	.4
Tryptophan	2.13
Ionizable	
Lysine	1.32
Arginine	3.8
Histidine	6.2
Aspartic acid	1.75
Glutamic acid	1.76
Asparagine	3.75
Glutamine	3.38
Proline	0
Cysteine	1.26

Table 36
Dissimilarities Based on Dipole Moment

		Gly																			
Ala,	A	0	A																		
Val,	V	0	0	V																	
Ile,	I	.1	.1	.1	I																
Leu,	L	0	0	0	.1	L															
Met,	M	0	0	0	.1	0	M														
Ser,	S	1.7	1.7	1.7	1.6	1.7	1.7	S													
Thr,	T	1.7	1.7	1.7	1.6	1.7	1.7	0	T												
Tyr,	Y	1.6	1.6	1.6	1.5	1.6	1.6	.1	.1	Y											
Phe,	F	.4	.4	.4	.3	.4	.4	1.3	1.3	1.2	F										
Trp,	W	.4	.4	.4	.3	.4	.4	1.3	1.3	1.2	0	W									
Lys,	K	1.3	1.3	1.3	1.2	1.3	1.3	.4	.4	.3	.9	.9	K								
His,	H	6.2	6.2	6.2	6.1	6.2	6.2	4.5	4.5	4.6	5.8	5.8	4.9	H							
Arg,	R	3.8	3.8	3.8	3.7	3.8	3.8	2.1	2.1	2.2	3.4	3.4	2.5	2.4	R						
Asp,	D	1.8	1.8	1.8	1.7	1.8	1.8	.1	.1	.2	1.4	1.4	.5	4.4	0	D					
Glu,	E	1.8	1.8	1.8	1.7	1.8	1.8	.1	.1	.2	1.4	1.4	.5	4.4	0	0	E				
Asn,	N	3.8	3.8	3.8	3.7	3.8	3.8	2.1	2.1	2.2	3.4	3.4	2.5	2.4	0	2.0	2.0	N			
Gln,	Q	3.4	3.4	3.4	3.3	3.4	3.4	1.7	1.7	1.8	3.0	3.0	2.1	2.8	1.6	1.6	1.6	.4	Q		
Pro,	P	0	0	0	.1	0	0	1.7	1.7	1.6	.4	.4	1.3	6.2	1.8	1.8	1.8	3.8	3.4	P	
Cys,	C	1.3	1.3	1.3	1.2	1.3	1.3	.4	.4	.3	.9	.9	0	4.9	.5	.5	.5	2.5	2.1	1.3	C

Table 37

Polarity Types of Amino Acid Replacement Pairs
in Human Hemoglobin

Polar - Polar	Frequency	
Glu -- Lys	12	n = 16
Asp -- His	2	
Arg -- His	2	
Non-polar - Non-polar		
Met -- Val	1	n = 2
Ser -- Cys	1	
Polar - Non-polar		
Gly -- Asp	8	n = 41
Glu -- Val	4	
Asp -- Asn	4	
Glu -- Gln	4	
His -- Tyr	4	
Glu -- Ala	3	
Lys -- Asn	3	
Gly -- Arg	3	
Glu -- Gly	2	
Arg -- Gln	2	
Arg -- Leu	2	
Lys -- Thr	1	
Asp -- Ala	1	

Using the model of the three dimensional structure of sperm whale myoglobin as typical of mammalian myoglobins provided by Perutz et al. (1965), locations of amino acid substitutions were related to exposure for sperm whale, harbor seal, and porpoise, see Tables 38, and 39. The polar amino acids seem to be favored as external substituents. In the sperm whale-harbor seal comparison, 15 out of the 24 substitutions are isopolar, and in the sperm whale-porpoise comparison, 7 out of 13 substitutions are isopolar. These results fit the suggested model better than the observations on hemoglobin polarity. A non-isopolar substitution may result in a "stabilizing" effect, i.e. polar-external, non-polar-in ternal. In the event that the effect is "non-stabilizing", a change in function may result. The "stabilizing" influence may, or may not result in an observable change in function.

The histidine points of heme attachment in hemoglobin and myoglobin provide another example of critical structure as do the cysteine points of heme attachment in cytochrome c. All three heme groups lie in non-polar "pockets".

One can argue then, that some amino acid replacements are more meaningful than others. However, it may very well be that at some places in a molecule, certain amino acids are equivalent to one another: glycine and alanine, serine and threonine, aspartic acid and glutamic acid, etc. If one accepts the principle that at non-critical points or region of the molecule, equivalent replacements can occur, then the more critical portions should be brought into focus readily. There are only seven points in the hemoglobin chain

Table 38

Polarity Types of Amino Acid Replacement Pairs
Sperm Whale vs. Harbor Seal Myoglobin

Sequence Position	AA Pair	Polar-Non-polar	I - E
1	Val - Gly	NP-NP	E
4	Glu - Asp	P-P	E
8	Asn - His	NP-P	E
12	His - Asn	P-NP	E
15	Ala - Gly	NP-NP	E
19	Ala - Thr	NP-P	E
21	Val - Leu	NP-NP	E
28	Ile - Val	NP-NP	I
45	Arg - Lys	P-P	E
51	Thr - Ser	P-P	I
53	Ala - Asp	NP-P	E
54	Glu - Asp	P-P	E
56	Lys - Arg	P-P	E
57	Ala - Arg	NP-P	E
66	Val - Asn	NP-NP	E
74	Ala - Gly	NP-NP	E
118	Arg - Lys	P-P	E
121	Gly - Ala	NP-NP	E
122	Asn - Glu	NP-P	E
129	Gly - Ala	NP-NP	E
132	Asn - Lys	NP-P	E
140	Lys - Asn	P-NP	E
151	Tyr - Phe	NP-NP	E
152	Asn - His	NP-P	E

I = Internal

E = External

Table 39

Polarity Types of Amino Acid Replacement Pairs
Sperm Whale vs. Porpoise Myoglobin

Sequence Position	AA Pair	Polar-Non-polar	I - E*
1	Val- Gly	NP-NP	E
12	His - Asn	P-NP	E
21	Val - Leu	NP-NP	E
28	Ile - Val	NP-NP	I
35	Ser - Gly	P-NP	I
66	Val - Asn	NP-NP	E
74	Ala - Gly	NP-NP	E
85	Glu - Asn	P-NP	E
121	Gly - Ala	NP-NP	E
122	Asn - Glu	NP-P	E
144	Ala - Thr	NP-P	E
151	Tyr - Phe	NP-NP	E
152	Asn - His	NP-P	E

*I = Internal
E = External

in the vertebrates that have not been seen to admit of replacements. This fortifies the basic contention that outside of the active site and volume of the molecule, there is a great deal of tolerance, so long as the general structure of the molecule is not disturbed. A certain amount of dimensional change is probably acceptable since the molecules are capable of some flexion.

Base Changes Required to Effect the Substitutions Seen

It is customary to consider the base changes required to effect the substitutions observed by assuming the change that occurs is the simplest one possible since base changes are subject to laws of probability. In keeping with the accepted approach, the simplest base changes capable of inducing the observed amino acid replacements are listed in (Table 40) with the understanding that there is no proof that they are, in fact, the actual changes.

Detailed Significance of Amino Acid Replacements of Fur Seal and Walrus Myoglobin

The replacements imposed by homology can be explained on the basis of the above principles with the assistance of some observations in the literature.

At positions #8 and #152, histidine is "replaced" by either glutamic acid or glutamine. Position #8 is externally directed and either residue would be capable of entering into hydrogen bonding with the surrounding water. A similar argument would hold for the replacement seen at position #152, since this is the penultimate residue at the carboxyl end of the molecule.

Table 40

Simplest Base Changes Required to Effect the Inferred Substitutions

Residue #	Amino Acid Pair	Base Change
8, 152	His - Glx	CAU - CAA (Gln) CAC - CAG (Gln) CAU - GAA (Glu) CAC - GAG (Glu)
19, 67	Thr - Ala	ACX - GCX (X = V, C, A, or G)
54	Asp - Glx	GAU - GAA (Glu) GAC - GAG (Glu) GAU - CAA (Gln) GAC - CAG (Gln)
56, 147	Arg - Lys	AGA - AAA AGG - AAG
66	Asn - Glx	AAU - CAA (Gln) AAC - CAG (Gln) AAU - GAA (Glu) AAC - GAG (Glu)
74	Glu - Gly	GAA - GGA GAG - GGG
76	Leu - Phe	CUU - UUU CUC - UUC UUA - UUU UUG - UUC
117	Ser - Gly	AGU - GGU AGC - GGC

At positions #19 and #67, there is a substitution of threonine for alanine. Position #19 is at the corner of the first bend in myoglobin. Alanine occupies this position in sperm whale myoglobin, and may very well be "buried in the crevices". (op cit) In the alpha and beta chains of horse and human hemoglobin, this position is occupied by glycine, glutamic acid, or aspartic acid. Position #67 is externally directed and one must invoke a "buried" alanine again.

Position #54 is the site of a substitution of glutamic acid or glutamine residue for an aspartic acid residue. Either replacement would maintain the hydrogen bonding quality at this position, one occurring at the corner between the fourth and fifth segments of the molecule.

At positions #56 and #147, arginine is replaced by lysine. The residue at #56 is externally directed at the same corner as #54. Position #147 is externally directed.

The substitution of a glutamic acid or glutamine residue for an asparagine at #66 is externally directed and either one would preserve the hydrogen bonding quality of the position.

The substitution of glycine for glutamic acid at position #74, an externally directed residue, would result in a "buried" hydrogen atom side chain. This would not enter into a hydrogen bond, nor would it interfere with the general external aqueous environment. In sperm whale myoglobin, this position is occupied by alanine.

The substitution of phenylalanine for leucine at position #76,

is probably trivial since both are non-polar and the residue is internally directed.

The substitution of glycine for serine at position #117 is non-disruptive since the externally directed glycine hydrogen will be "buried".

It is possible, then to treat the persistence of the observed amino acid differences on the basis of the proposed model.

Most of the differences between the myoglobin sequences of the phocidae and the walrus-fur seal myoglobin are readily explained as "acceptable" on the basis of hydrogen bonding and hydrophobic bonding characteristics. The appearance of glycine or alanine as an externally directed residue is explained on the grounds that the small side chains involved are buried in the crevices of the molecule.

CONCLUSIONS

I. This research was designed to answer the question "What is the degree of difference or similarity seen in the primary structure of the myoglobins of four members of the family Phocidae, and what is the degree of difference or similarity seen in the primary structure of the myoglobins of members of the three families of the Pinnipedia?" A detailed response can now be given. This response is made with the implicit understanding that the inferred sequences are based on the two homology assumptions previously stated.

1. The primary structure of the myoglobins of harbor seals from the Bering Sea and the coast of Maine are probably identical, with the possible exception of those portions of the sequence represented by the peptides T-1e and T-12b for which no information was obtained.

2. The primary structure of the myoglobins of harbor seal and ribbon seal are probably identical, with the possible exceptions of those portions of the sequence represented by peptides T-1e and T-12a.

3. The primary structure of myoglobins of the bearded seal and Weddell seal are both probably identical with that of harbor seal with the possible exception of those portions of the sequence represented by peptides T-1e, T-12b, and T-13 for which no information was obtained.

4. The primary structure of the myoglobin of the northern fur seal possesses at least eight replacement differences when compared with that of the harbor seal. This comparison does not include those portions of the sequence represented by peptides T-1e, T-4, T-8, T-10c, T-12b, and T-13 for which no information was obtained.

6. The primary structure of the myoglobin of the walrus differs from that of the fur seal by one amino acid replacement. This comparison is based on those portions of the sequence that were found for the walrus myoglobin.

7. With the above exceptions noted, there are no differences in the myoglobins of four members of the family Phocidae. There are at least nine differences between the myoglobins of family Phocidae and family Odobenidae. There are at least eight differences between the myoglobins of family Phocidae and family Otariidae. There is at least one difference between the myoglobins of family Odobenus and family Otariidae.

II. The peptide mapping method employed for this research has definite potential for a study of this type. The convenience of handling small pieces of paper, the short time required for electrophoresis, the minimal support space required, and relatively inexpensive apparatus are all definite advantages. As previously stated, a solid support medium other than cellulose paper would enhance this technique immeasurably. Even so, this present method would permit qualitative surveys to be carried out quickly and inexpensively. Such surveys are necessary to the proper planning of use of such formidable tools as the sequenator.

LITERATURE CITED

Bieleski, R. L. 1965. Separation of phosphate esters by thin layer chromatrography and electrophoresis. Anal. Biochem. 12:230-234.

Black, J. A. and G. Leaf. 1965. The action of cyanogen bromide on horse-heart cytochrome c and horse heart myoglobin. Biochem. J. 96: 693-699.

Blackburn, S. 1968. Amino acid determination. Marcel Dekker, Inc. New York. pp. 271.

Boyer, S. H., E. F. Crosby, T. F. Thurman, G. F. Fuller, S. E. Leslie, M. K. Shepard, and C. N. Herndon. 1969. Hemoglobins A and A_2 in new world primates: Comparative variation and its evolutionary implications. Science 166:1428-1431.

Boyer, S. H., D. C. Fainer, and M. A. Naughton. 1963. Myoglobin: Inherited structural variation in man. Science 140:1228-1231.

Bennett, J. C. 1967. Paper chromatography and electrophoresis: Special procedure for peptide maps. Method of Enzymology 11:330-338.

Bradshaw, R. A., W. H. Ganer, and F. R. N. Gurd. 1969. Comparison of myoglobins from harbor seal, porpoise, and sperm whale. III. Isolation and characterization of the tryptic peptides of harbor seal myoglobin. J. Biol. Chem. 244:2;49-2158.

Bradshaw, R. A., F. R. N. Gurd. 1969. Comparison of myoglobins from harbor seal, porpoise, and sperm whale. V. The complete amino acid sequence of harbor seal and porpoise myoglobins. J. Biol. Chem. 244:2167-2181.

Burns, D. G. W., and N. A. Turner. 1967. Peptide mapping on cellulose thin layers. J. Chromatog. 30:469-475.

Burns, J. J. and F. H. Fay. In Press. Comparative morphology of the skull of the ribbon seal, Histriophoca fasciata (Zimmerman), with remarks on systematics of Phocidae. J. Zool.

Dautrevaux, M., K. Han, X. Chaila, and G. Biserte. 1969. Sequence partielle des acides amines de la myoglobine de boeuf. Bull. Soc. Chim. Biol. 51:989-1000.

Dayhoff, M. O. 1969. Atlas of Protein Sequence and Structure. Vol. 4. National Biochemical Research Foundation. Silver Spring, Maryland pp. 350.

Dickerson, R. E. and I. Geis. 1969. The structure and action of proteins. Harper and Row. New York, N.Y. pp. 119.

Edman, P., and G. Begg. 1967. A protein sequenator. European J. Biochem. 1:80-91.

Edmundson, A. B. 1963. Cleavage of sperm whale myoglobin with cyanogen bromide. Nature 196:354-357.

Edmundson, A. B. 1965. Amino acid sequence of sperm whale myoglobin. Nature 205:883-887.

Edmundson, A. B., and C. H. W. Hirs. 1962a. On the structure of sperm whale myoglobin. I. The amino acid composition and termination groups of the chromatographically purified protein. J. Mol. Biol. 5:663-682.

Edmundson, A. B., and C. H. W. Hirs. 1962b. On the structure of sperm whale myoglobin. II. The tryptic hydrolysis of the denatured protein. J. Mol. Biol. 5:683-705.

Edmundson, A. B., and C. H. W. Hirs. 1962c. On the structure of sperm whale myoglobin. III. Amino acid sequences of the smaller tryptic peptides containing aromatic residues. J. Mol. Biol. 5:706-708.

Fay, F. H., V. R. Rausch, and E. T. Feltz. 1967. Cytogenetic comparisons of some pinnipeds (Mammalia:Eutheria). Can. J. Zool. 45:773-778.

Fitch, W. M. 1967. Evidence suggesting a non-random character to nucleotide replacements in naturally occurring mutations. J. Mol. Biol. 20: 499-507.

Genaux, C. 1969. The comparative amino acid composition of five hemoglobins from the genus Microtus Ph.D. dissertation, University of Alaska, College, Alaska.

Gross, E. and B. Witkop. 1962. Nonenzymatic cleavage of peptide bonds: The methionine residues in bovine pancreatic ribonuclease. J. Biol. Chem. 237:1856-1860.

Hamilton, P. B. 1963. Ion exchange chromatography of amino acids: A single column, high resolving a fully automatic procedure. Anal. Chem. 35: 2055-2964.

Hamilton, P. B., and R. Anderson. 1959. Ion exchange chromatography of amino acids: Semiautomatic method of operation with cationic exchange resin columns. Anal. Chem. 31:1504-1512.

Hill, R. L. and C. M. Harris, J. F. Naylor, and W. M. Sams. 1969. The partial amino acid sequence of human myoglobin. J. Biol. Chem. 244:2182-2194.

Hirs, C. H. W., S. Moore, and W. H. Stein. 1960. The sequence of the amino acid residues in performic acid-oxidized ribonuclease. J. Biol. Chem. 235:633-647.

Hoerman, K. C. and K. Kamel. 1967. Recycled chromatograms for better peptide mapping. Anal. Biochem. 21:107-110.

Ingram, V. M. 1958. Abnormal human hemoglobins. I. The comparison of normal and sickle-cell hemoglobin by "fingerprinting". Biochim. Biophys. Acta. 28:539-545.

James, N. T. 1968. Histochemical demonstration of myoglobin in skeletal muscle fibers and muscle spindles. Nature 219:1174-1175.

Jegorov, C., B. Keil, and F. Sorm. 1965. Comparative study of the primary structure of myoglobins of various animals by the fingerprinting method. Coll. Czechoslov. Chem. Communic. 30:105-117.

King, J. L. and T. H. Jukes. 1969. Non-Darwinian evolution. Science 164: 788-798.

Margoliash, E. 1969. Homology: A definition. Science 163:127.

Margoliash, E. and A. Schejter. 1966. Cytochrome c bonds: Constancy of residue types. Advanc. Protein Chem. 21:113-286.

McClellan, A. L. 1963. Tables of experimental dipole moments. W. H. Freeman Co., San Francisco, Calif. pp. 713.

Mefford, R. B., Jr., R. M. Summers, and J. D. Clayton. 1968. Correlational analysis of seven procedures for determining area of chromatographic curves. J. Chromatog. 35:469-474.

Mondino, A. 1967. A new system of automatic amino acid analysis, Part I. J. Chromatog. 30:110-112.

Moore, S., D. H. Spackman, and W. H. Stein. 1958. Chromatography of amino acids on sulfonated polystyrene resins. Anal. Chem. 30:1185-1190.

Moore, S. and W. H. Stein. 1951. Chromatography of amino acids on sulfonated polystyrene resins. J. Biol. Chem. 192:663-681.

Nolan, C. and E. Margoliash. 1968. Comparative aspects of primary structures of proteins. Ann. Rev. of Biochem. pp. 37.

Perutz, M. F. 1965. Structure and function of hemoglobin. I. A tentative atomic model of horse oxyhaemoglobin. J. Mol. Biol. 13:646-668.

Perutz, M. F., J. C. Kendrew, and H. C. Watson. 1965. Structure and function of haemoglobin: Some relation between polypeptide chain configuration and amino acid sequence. J. Mol. Biol. 13:669-678.

Peterson, R. F., L. W. Nauman, and D. F. Hamilton. 1966. Amino acid composition of six distinct types of beta casein. J. Dairy Sci. 49: 601-607.

Piez, K. A. and L. Morris. 1960. A modified procedure for the automatic analysis of the amino acids. Anal. Biochem. 1:187-201.

Richmond, R. C. 1970. Non-Darwinian Evolution: A critique. Nature 225: 1025-1028.

Romer, A. S. 1962. Vertebrate Paleontology. University of Chicago Press, Chicago, Ill. pp. 468.

Rovery, M. 1967. Chromatography of trypsin to remove chymotrypsin and of chymotrypsin to remove trypsin. Methods of Enzymology Vol. XI. pp. 231-236. Academic Press, N.Y.

Sarich, V. M. 1969a. Pinniped origins and the rate of evolution of carnivore albumins. Syst. Zool. 18:286-295.

Sarich, V. M. 1969b. Pinniped phylogeny. System. Zool. 18:416-422.

Schaffer, N. K., R. A. Balakir, V. Riley, and T. E. Moffitt. 1966. O-Ring joint tubes for protein hydrolyzates. Anal. Biochem. 15:93-96.

Scheffer, V. B. 1968. Seals, Sea Lions, and Walruses. Stanford University Press, Stanford, Calif. pp. 179.

Schroeder, W. A. 1968. The primary structure of proteins. Harper and Row. New York, N. Y. pp. 210.

Schroeder, W. A., W. R. Helmquist, and J. Roger Shelton. 1966. A nomogram for calculating the results of an automatic amino acid analysis. Anal. Chem. 38:1281-1285.

Smyth, D. G., W. H. Stein, and S. Moore. 1963. The sequence of amino acid residues in bovine pancreatic ribonuclease: Revisions and confirmations. J. Biol. Chem. 238:227.

Sneath, P. H. A. 1965. Relations between chemical structure and biological activity in proteins. J. Theoretical. Biol. 12: 157-195.

Snell, E. E. 1945. The microbiological assay of amino acids. Advances In Protein Chemistry. Vol. II. pp. 85-118.

Sochard, M. 1968. Abnormal human hemoglobins. Atlas of Protein Structure and Sequence. pp. 47-56. National Biochem. Research Foundation, Silver Spring, Md.

Sokol, R. R. and P. H. A. Sneath. 1963. Principles of Numerical Taxonomy. W. H. Freeman and Co., San Francisco, Calif.

Spackman, D. H., W. H. Stein, and S. Moore. 1958. Automatic recording apparatus for use in the chromatography of amino acids. Anal. Chem. 30:1190-1210.

Stockell, A. 1961. Comparative studies on tryptic digests of myoglobins. J. Mol. Biol. 3:362-366.

Thompson, R. A. and B. J. Miles. 1964. Ion-exchange chromatography: Improvements in the single column system. Nature 203:483-484.

Tristram, G. R. 1949. Amino acid composition of purified proteins. Advances in Protein Chemistry. 5:84-153.

Vogel, F. 1969. Point mutations and human hemoglobin variants. Humangenetik 8:1-26.

Wang, A. and H. E. Sutton. 1965. Human transferrins C & D: Chemical difference in a peptide. Science 149:435-437.